Diffusion in a weakly random Hamiltonian flow

Tomasz Komorowski * Lenya Ryzhik †

February 7, 2008

Abstract

We consider the motion of a particle governed by a weakly random Hamiltonian flow. We identify temporal and spatial scales on which the particle trajectory converges to a spatial Brownian motion. The main technical issue in the proof is to obtain error estimates for the convergence of the solution of the stochastic acceleration problem to a momentum diffusion. We also apply our results to the system of random geometric acoustics equations and show that the energy density of the acoustic waves undergoes a spatial diffusion.

1 Introduction

The long time, large distance behavior of a massive particle in a weakly random time-independent potential field is described by the momentum diffusion: the particle momentum undergoes the Brownian motion on the energy sphere. This intuitive result has been first proved in [9] in dimensions higher than two, and later extended to two dimensions with the Poisson distribution of scatterers in [2]. On the other hand, the long time limit of a momentum diffusion is the standard spatial Brownian motion. Hence, a natural question arises if it is possible to obtain such a Brownian motion directly as the limiting description in the original problem of a particle in a quenched random potential. This necessitates the control of the particle behavior over times longer than those when the momentum diffusion holds.

We consider a particle that moves in an isotropic weakly random Hamiltonian flow with the Hamiltonian of the form $H_\delta(\mathbf{x},\mathbf{k}) = H_0(k) + \sqrt{\delta}H_1(\mathbf{x},k)$, $k = |\mathbf{k}|$, and $\mathbf{x},\mathbf{k} \in \mathbb{R}^d$ with $d \geq 3$:

$$\frac{dX^\delta}{dt} = \nabla_{\mathbf{k}} H_\delta, \quad \frac{dK^\delta}{dt} = -\nabla_{\mathbf{x}} H_\delta, \quad X^\delta(0) = 0, \quad K^\delta(0) = \mathbf{k}_0. \tag{1.1}$$

Here $H_0(k)$ is the background Hamiltonian and $H_1(\mathbf{x},k)$ is a random perturbation. As we have mentioned, it has been shown in [9] that, when $H_\delta(\mathbf{x},\mathbf{k}) = \dfrac{k^2}{2} + \sqrt{\delta}V(\mathbf{x})$, and under certain mixing assumptions on the random potential $V(\mathbf{x})$, the momentum process $K^\delta(t/\delta)$ converges to a diffusion process $K(t)$ on the sphere $k = k_0$ and the rescaled spatial component $\tilde{X}^\delta(t) = \delta X^\delta(t/\delta)$ converges to $X(t) = \int_0^t K(s)ds$. This is the momentum diffusion mentioned above. Another special case,

$$H_\delta(\mathbf{x},\mathbf{k}) = (c_0 + \sqrt{\delta}c_1(\mathbf{x}))|\mathbf{k}|, \tag{1.2}$$

*Institute of Mathematics, UMCS, pl. Marii Curie Skłodowskiej 1, 20-031 Lublin, Poland; e-mail: komorow@hektor.umcs.lublin.pl

†Department of Mathematics, University of Chicago, Chicago, IL 60637, USA; e-mail: ryzhik@math.uchicago.edu

arises in the geometrical optics limit of wave propagation. Here c_0 is the background sound speed, and $c_1(\mathbf{x})$ is a random perturbation. This case has been considered in [1], where it has been shown that, once again, $K^\delta(t/\delta)$ converges to a diffusion process $K(t)$ on the sphere $\{k = k_0\}$ while $\tilde{X}^\delta(t) = \delta X^\delta(t/\delta)$ converges to $X(t) = c_0 \int_0^t \hat{K}(s)ds$, $\hat{K}(t) := K(t)/|K(t)|$.

We show in this paper that this analysis may be pushed beyond the time of the momentum diffusion, and that under certain assumptions concerning the mixing properties of H_1 in the spatial variable there exists $\alpha_0 > 0$ so that the process $\delta^{1+\alpha} X^\delta(t/\delta^{1+2\alpha})$ converges to the standard Brownian motion in $\mathbb{R}^d$ for all $\alpha \in (0, \alpha_0)$. The main difficulty of the proof is to obtain error estimates in the convergence of $K^\delta(\cdot)$ to the momentum diffusion on time scales of the order δ^{-1}. The error estimates allow us to push the analysis to times much longer than δ^{-1} where the momentum diffusion converges to the standard Brownian motion. The method of the proof is a modification of the cut-off technique used in [1] and [9].

A similar question arises in the semi-classical limit of the quantum mechanics and high frequency wave propagation. The Wigner transform [5], or the phase space energy density of the solution of the Schrödinger equation, is approximated in a weakly random medium by the solution of a deterministic linear Boltzmann equation [3]. This behavior is also conjectured for the acoustic and other waves in a weakly random medium [12]. As in the momentum diffusion model for a particle, the long time limit of the Boltzmann equation is the spatial diffusion equation. It has been recently shown in [4] that, indeed, one may push the analysis of [3] beyond the times on which the Boltzmann equation holds and obtain the diffusive behavior of the energy density of the solutions of the Schrödinger equation in the weak coupling limit.

We also apply our results to the problem of multiple scattering of the acoustic waves. Our approach is different from that of [4] mentioned above: we first consider the random geometrical optics approximation of the wave phase space energy density. The rays in the phase space satisfy the Hamiltonian equations (1.1) with the Hamiltonian given by (1.2). Therefore, the aforementioned convergence result of the solutions of (1.1) to the standard Brownian motion, combined with the error estimates on the geometrical optics approximation of the Wigner distribution of the solutions of the wave equation, allows us to establish rigorously the diffusive behavior of the wave energy density. To the best of our knowledge, this is the first result of such kind for classical waves.

This paper is organized as follows. Section 2 contains the main results on the convergence of solutions of (1.1) to the Brownian motion as well as the assumptions on the random medium. Sections 3 and 4 contain the proof of our main results, Theorems 2.1, 2.5 and 2.7: the error estimates on the passage to the momentum diffusion and the passage from the momentum diffusion to the spatial diffusion, respectively. Section 5 discusses the application to the wave equation. Finally, Appendix A contains the proof of a technical result that is stated in Lemma 3.5. We note that all constants appearing throughout the paper do not depend on $\delta \in (0, 1]$ unless otherwise specified.

Acknowledgment. The research of TK was partially supported by KBN grant 2PO3A03123. The work of LR was partially supported by an NSF grant DMS-0203537, an ONR grant N00014-02-1-0089 and an Alfred P. Sloan Fellowship.

2 The main result and preliminaries

2.1 The notation

As we will avoid the singular point $\mathbf{k} = 0$, we denote $\mathbb{R}^d_* := \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $\mathbb{R}^{2d}_* := \mathbb{R}^d \times \mathbb{R}^d_*$. Also $\mathbb{S}^{d-1}_R(\mathbf{x})$ ($\mathbb{B}_R(\mathbf{x})$) shall stand for a sphere (open ball) in $\mathbb{R}^d$ of radius $R > 0$ centered at $\mathbf{x}$. We shall drop writing either $\mathbf{x}$, or R in the notation of the sphere (ball) in the particular cases when either $\mathbf{x} = 0$, or $R = 1$. For a fixed $M > 0$ we define the spherical shell $A(M) := [\mathbf{k} \in \mathbb{R}^d_* : M^{-1} \leq |\mathbf{k}| \leq M]$

in the $\mathbf{k}$-space, and $\mathcal{A}(M) := \mathbb{R}^d \times A(M)$ in the whole phase space. Given a vector $\mathbf{v} \in \mathbb{R}^d_*$ we denote by $\hat{\mathbf{v}} := \mathbf{v}/|\mathbf{v}| \in \mathbb{S}^{d-1}$ the unit vector in the direction of $\mathbf{v}$. For any set A we shall denote by A^c its complement.

For any non-negative integers p, q, r, positive times $T > T_* \geq 0$ and a function $G : [T_*, T] \times \mathbb{R}^{2d}_* \to \mathbb{R}$ that has p, q and r derivatives in the respective variables we define

$$\|G\|^{[T_*,T]}_{p,q,r} := \sum \sup_{(t,\mathbf{x},\mathbf{k}) \in [T_*,T] \times \mathbb{R}^{2d}} |\partial^\alpha_t \partial^\beta_{\mathbf{x}} \partial^\gamma_{\mathbf{k}} G(t, \mathbf{x}, \mathbf{k})|. \tag{2.1}$$

The summation range covers all integers $0 \leq \alpha \leq p$ and all integer valued multi-indices $|\beta| \leq q$ and $|\gamma| \leq r$. In the special case when $T_* = 0$, $T = +\infty$ we write $\|G\|_{p,q,r} = \|G\|^{[0,+\infty)}_{p,q,r}$. We denote by $C^{p,q,r}_b([0, +\infty) \times \mathbb{R}^{2d}_*)$ the space of all functions G with $\|G\|_{p,q,r} < +\infty$. We shall also consider spaces of bounded and a suitable number of times continuously differentiable functions $C^{p,q}_b(\mathbb{R}^{2d}_*)$ and $C^p_b(\mathbb{R}^d_*)$ with the respective norms $\| \cdot \|_{p,q}$ and $\| \cdot \|_p$.

2.2 The background Hamiltonian

We assume that the background Hamiltonian $H_0(k)$ is isotropic, that is, it depends only on $k = |\mathbf{k}|$, and is uniform in space. Moreover, we assume that $H_0 : [0, +\infty) \to \mathbb{R}$ is a strictly increasing function satisfying $H_0(0) \geq 0$ and such that it is of C^3-class of regularity in $(0, +\infty)$ with $H_0'(k) > 0$ for all $k > 0$, and let

$$h^*(M) := \max_{k \in [M^{-1},M]} (H_0'(k) + |H_0''(k)| + |H_0'''(k)|), \quad h_*(M) := \min_{k \in [M^{-1},M]} H_0'(k). \tag{2.2}$$

Two examples of such Hamiltonians are the quantum Hamiltonian $H_0(k) = k^2/2$ and the acoustic wave Hamiltonian $H_0(k) = c_0 k$.

2.3 The random medium

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, and let $\mathbb{E}$ denote the expectation with respect to $\mathbb{P}$. We denote by $\|X\|_{L^p(\Omega)}$ the L^p-norm of a given random variable $X : \Omega \to \mathbb{R}$, $p \in [1, +\infty]$. Let $H_1 : \mathbb{R}^d \times [0, +\infty) \times \Omega \to \mathbb{R}$ be a random field that is measurable and strictly stationary in the first variable. This means that for any shift $\mathbf{x} \in \mathbb{R}^d$, $k \in [0, +\infty)$, and a collection of points $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ the laws of $(H_1(\mathbf{x}_1 + \mathbf{x}, k), \ldots, H_1(\mathbf{x}_n + \mathbf{x}, k))$ and $(H_1(\mathbf{x}_1, k), \ldots, H_1(\mathbf{x}_n, k))$ are identical. In addition, we assume that $\mathbb{E}H_1(\mathbf{x}, k) = 0$ for all $k \geq 0$, $\mathbf{x} \in \mathbb{R}^d$, the realizations of $H_1(\mathbf{x}, k)$ are $\mathbb{P}$–a.s. C^2-smooth in $(\mathbf{x}, k) \in \mathbb{R}^d \times (0, +\infty)$ and they satisfy

$$D_{i,j}(M) := \max_{|\alpha|=i} \operatorname*{ess-sup}_{(\mathbf{x},k,\omega) \in \mathbb{R}^d \times [M^{-1},M] \times \Omega} |\partial^\alpha_{\mathbf{x}} \partial^j_k H_1(\mathbf{x}, k; \omega)| < +\infty, \quad i, j = 0, 1, 2. \tag{2.3}$$

We define $\tilde{D}(M) := \sum_{0 \leq i+j \leq 2} D_{i,j}(M)$.

We suppose further that the random field is strongly mixing in the uniform sense. More precisely, for any $R > 0$ we let $\mathcal{C}^i_R$ and $\mathcal{C}^e_R$ be the σ–algebras generated by random variables $H_1(\mathbf{x}, k)$ with $k \in [0, +\infty)$, $\mathbf{x} \in \mathbb{B}_R$ and $\mathbf{x} \in \mathbb{B}^c_R$ respectively. The uniform mixing coefficient between the σ–algebras is

$$\phi(\rho) := \sup[\, |\mathbb{P}(B) - \mathbb{P}(B|A)| : \ R > 0, \ A \in \mathcal{C}^i_R, \ B \in \mathcal{C}^e_{R+\rho}\,],$$

for all $\rho > 0$. We suppose that $\phi(\rho)$ decays faster than any power: for each $p > 0$

$$h_p := \sup_{\rho \geq 0} \rho^p \phi(\rho) < +\infty. \tag{2.4}$$

The two-point spatial correlation function of the random field H_1 is $R(\mathbf{y},k) := \mathbb{E}[H_1(\mathbf{y},k)H_1(\mathbf{0},k)]$. Note that (2.4) implies that for each $p > 0$

$$h_p(M) := \sum_{i=0}^{4} \sum_{|\alpha|=i} \sup_{(\mathbf{y},k)\in\mathbb{R}^d\times[M^{-1},M]} (1+|\mathbf{y}|^2)^{p/2}|\partial_{\mathbf{y}}^{\alpha} R(\mathbf{y},k)| < +\infty, \quad M > 0. \tag{2.5}$$

We also assume that the correlation function $R(\mathbf{y},l)$ is of the C^∞-class for a fixed $l > 0$, is sufficiently smooth in l, and that for any fixed $l > 0$

$$\hat{R}(\mathbf{k},l) \text{ does not vanish identically on any hyperplane } H_{\mathbf{p}} = \{\mathbf{k} : (\mathbf{k}\cdot\mathbf{p}) = 0\}. \tag{2.6}$$

Here $\hat{R}(\mathbf{k},l) = \int R(\mathbf{x},l)\exp(-i\mathbf{k}\cdot\mathbf{x})d\mathbf{x}$ is the power spectrum of H_1.

The above assumptions are satisfied, for example, if $H_1(\mathbf{x},k) = c_1(\mathbf{x})h(k)$, where $c_1(\mathbf{x})$ is a stationary uniformly mixing random field with a smooth correlation function, and $h(k)$ is a smooth deterministic function.

2.4 Certain path-spaces

For fixed integers $d, m \geq 1$ we let $\mathcal{C}^{d,m} := C([0,+\infty); \mathbb{R}^d \times \mathbb{R}^m_*)$: we shall omit the subscripts in the notation of the path space if $m = d$. We define $(X(t), K(t)) : \mathcal{C}^{d,m} \to \mathbb{R}^d \times \mathbb{R}^m_*$ as the canonical mapping $(X(t;\pi), K(t;\pi)) := \pi(t)$, $\pi \in \mathcal{C}^{d,m}$ and also let $\theta_s(\pi)(\cdot) := \pi(\cdot + s)$ be the standard shift transformation.

For any $u \leq v$ denote by $\mathcal{M}_u^v$ the σ-algebra of subsets of $\mathcal{C}$ generated by $(X(t), K(t))$, $t \in [u,v]$. We write $\mathcal{M}^v := \mathcal{M}_0^v$ and $\mathcal{M}$ for the σ algebra of Borel subsets of $\mathcal{C}$. It coincides with the smallest σ–algebra that contains all $\mathcal{M}^t$, $t \geq 0$.

Let $\delta_*(M) := H_0\left(M^{-1}\right)/(2\tilde{D}(M))$. For a given $M > 0$ and $\delta \in (0, \delta_*(M)]$ we let

$$M_\delta := \max\left\{ H_0^{-1}(H_0(M) + 2\sqrt{\delta}\tilde{D}(M)), \left[H_0^{-1}\left(H_0\left(\frac{1}{M}\right) - 2\sqrt{\delta}\tilde{D}(M)\right)\right]^{-1} \right\}. \tag{2.7}$$

For a particle that is governed by the Hamiltonian flow generated by $H_\delta(\mathbf{x},\mathbf{k})$ we have $M_\delta^{-1} \leq |K(t)| \leq M_\delta$ for all t provided that $K(0) \in A(M)$. Accordingly, we define $\mathcal{C}(T,\delta)$ as the set of paths $\pi \in \mathcal{C}$ so that both $(2M_\delta)^{-1} \leq |K(t)| \leq 2M_\delta$, and

$$\left| X(t) - X(u) - \int_u^t H_0'(K(s))\hat{K}(s)ds \right| \leq \tilde{D}(2M_\delta)\sqrt{\delta}(t-u), \text{ for all } 0 \leq u < t \leq T.$$

In the case when $\delta = 1$, or $T = +\infty$ we shall write simply $\mathcal{C}(T)$, or $\mathcal{C}(\delta)$ respectively.

2.5 The main results

Let the function $\phi_\delta(t,\mathbf{x},\mathbf{k})$ satisfy the Liouville equation

$$\begin{aligned} &\frac{\partial\phi^\delta}{\partial t} + \nabla_{\mathbf{x}} H_\delta(\mathbf{x},\mathbf{k})\cdot\nabla_{\mathbf{k}}\phi^\delta - \nabla_{\mathbf{k}} H_\delta(\mathbf{x},\mathbf{k})\cdot\nabla_{\mathbf{x}}\phi^\delta = 0, \\ &\phi^\delta(0,\mathbf{x},\mathbf{k}) = \phi_0(\delta\mathbf{x},\mathbf{k}). \end{aligned} \tag{2.8}$$

We assume that the initial data $\phi_0(\mathbf{x},\mathbf{k})$ is a compactly supported function four times differentiable in $\mathbf{k}$, twice differentiable in $\mathbf{x}$ whose support is contained inside a spherical shell $\mathcal{A}(M) = \{(\mathbf{x},\mathbf{k}) : M^{-1} < |\mathbf{k}| < M\}$ for some positive $M > 0$.

Let us define the diffusion matrix D_{mn} by

$$D_{mn}(\hat{\mathbf{k}},l) = -\frac{1}{2}\int_{-\infty}^{\infty} \frac{\partial^2 R(H_0'(l)s\hat{\mathbf{k}},l)}{\partial x_n \partial x_m} ds = -\frac{1}{2H_0'(l)}\int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{\mathbf{k}},l)}{\partial x_n \partial x_m} ds, \quad m,n=1,\ldots,d. \tag{2.9}$$

Then we have the following result.

Theorem 2.1 *Let ϕ^δ be the solution of (2.8) and let $\bar{\phi}$ satisfy*

$$\begin{aligned}\frac{\partial \bar{\phi}}{\partial t} &= \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial \bar{\phi}}{\partial k_n}\right) + H_0'(k)\,\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}\bar{\phi} \\ \bar{\phi}(0,\mathbf{x},\mathbf{k}) &= \phi_0(\mathbf{x},\mathbf{k}).\end{aligned} \tag{2.10}$$

Suppose that $M \geq M_0 > 0$ and $T \geq T_0 > 0$. Then, there exist two constants C, $\alpha_0 > 0$ such that for all $T \geq T_0$

$$\sup_{(t,\mathbf{x},\mathbf{k})\in[0,T]\times K}\left|\mathbb{E}\phi^\delta\left(\frac{t}{\delta},\frac{\mathbf{x}}{\delta},\mathbf{k}\right) - \bar{\phi}(t,\mathbf{x},\mathbf{k})\right| \leq CT(1+\|\phi_0\|_{1,4})\delta^{\alpha_0} \tag{2.11}$$

for all compact sets $K \subset \mathcal{A}(M)$.

Remark 2.2 We shall denote by C, $C_1,\ldots$, α_0, $\alpha_1,\ldots$, γ_0, $\gamma_1,\ldots$ throughout this article generic positive constants. Unless specified otherwise the constants denoted this way *shall depend neither on δ, nor on T.* We will also assume that $T \geq T_0 > 0$ and $M \geq M_0 > 0$.

Remark 2.3 Classical results of the theory of stochastic differential equations, see e.g. Theorem 6 of Chapter 2, p. 176 and Corollary 4 of Chapter 3, p. 303 of [6], imply that there exists a unique solution to the Cauchy problem (2.10) that belongs to the class $C_b^{1,1,2}([0,+\infty)\times\mathbb{R}_*^{2d})$. This solution admits a probabilistic representation using the law of a time homogeneous diffusion $\mathfrak{Q}_{\mathbf{x},\mathbf{k}}$ whose Kolmogorov equation is given by (2.10), see Section 3.6 below.

Note that

$$\begin{aligned}\sum_{m=1}^{d} D_{nm}(\hat{\mathbf{k}},k)\hat{k}_m &= -\sum_{m=1}^{d}\frac{1}{2H_0'(k)}\int_{-\infty}^{\infty}\frac{\partial^2 R(s\hat{\mathbf{k}},k)}{\partial x_n\partial x_m}\hat{k}_m ds \\ &= -\sum_{m=1}^{d}\frac{1}{2H_0'(k)}\int_{-\infty}^{\infty}\frac{d}{ds}\left(\frac{\partial R(s\hat{\mathbf{k}},k)}{\partial x_n}\right)ds = 0\end{aligned}$$

and thus the K-process generated by (2.10) is indeed a diffusion process on a sphere $k = \text{const}$, or, equivalently, equations (2.10) for different values of k are decoupled. Assumption (2.6) implies the following.

Proposition 2.4 *The matrix $D(\hat{\mathbf{k}},l)$ has rank $d-1$ for each $\hat{\mathbf{k}} \in \mathbb{S}^{d-1}$ and each $l > 0$.*

The proof is the same as that of Proposition 4.3 in [1]. It can be shown, using the argument given on pp. 122-123 of ibid., that, under assumption (2.6), equation (2.10) is hypoelliptic on the manifold $\mathbb{R}^d \times \mathbb{S}_k^{d-1}$ for each $k > 0$.

We also show that solutions of (2.10) converge in the long time limit to the solutions of the spatial diffusion equation. More, precisely, we have the following result. Let $\bar{\phi}_\gamma(t,\mathbf{x},\mathbf{k}) = \bar{\phi}(t/\gamma^2,\mathbf{x}/\gamma,\mathbf{k})$,

where $\bar{\phi}$ satisfies (2.10) with an initial data $\bar{\phi}_\gamma(0,t,\mathbf{x},\mathbf{k}) = \phi_0(\gamma\mathbf{x},\mathbf{k})$. We also let $w(t,\mathbf{x},k)$ be the solution of the spatial diffusion equation:

$$\frac{\partial w}{\partial t} = \sum_{m,n=1}^{d} a_{mn}(k)\frac{\partial^2 w}{\partial x_n \partial x_m}, \qquad (2.12)$$
$$w(0,\mathbf{x},k) = \bar{\phi}_0(\mathbf{x},k)$$

with the averaged initial data

$$\bar{\phi}_0(\mathbf{x},k) = \frac{1}{\Gamma_{d-1}}\int_{\mathbb{S}^{d-1}} \phi_0(\mathbf{x},\mathbf{k})d\Omega(\hat{\mathbf{k}}).$$

Here $d\Omega(\hat{\mathbf{k}})$ is the surface measure on the unit sphere $\mathbb{S}^{d-1}$ and Γ_n is the area of an n-dimensional sphere. The diffusion matrix $A := [a_{nm}]$ in (2.12) is given explicitly as

$$a_{nm}(k) = \frac{1}{\Gamma_{d-1}}\int_{\mathbb{S}^{d-1}} H_0'(k)\hat{k}_n\chi_m(\mathbf{k})d\Omega(\hat{\mathbf{k}}). \qquad (2.13)$$

The functions χ_j appearing above are the mean-zero solutions of

$$\sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial \chi_j}{\partial k_n}\right) = -H_0'(k)\hat{k}_j. \qquad (2.14)$$

Note that equations (2.14) for χ_m are elliptic on each sphere $\{|\mathbf{k}| = k\}$. This follows from the fact that the equations for each such sphere are all decoupled and Proposition 2.4. Also note that the matrix A is positive definite. Indeed, let $\mathbf{c} = (c_1,\ldots,c_d) \in \mathbb{R}^d$ be a fixed vector and let $\chi_{\mathbf{c}} := \sum_{m=1}^{d} c_m\chi_m$. Since the matrix D is non-negative we have

$$\begin{aligned}(A\mathbf{c},\mathbf{c})_{\mathbb{R}^d} &= -\frac{1}{\Gamma_{d-1}}\sum_{m,n=1}^{d}\int_{\mathbb{S}^{d-1}} \chi_{\mathbf{c}}(\hat{\mathbf{k}},l)\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},l)\frac{\partial \chi_{\mathbf{c}}(\hat{\mathbf{k}},l)}{\partial k_n}\right)d\Omega(\hat{\mathbf{k}}) \qquad (2.15)\\ &= -\frac{1}{\Gamma_{d-1}}\sum_{m,n=1}^{d}\int_{\mathbb{R}^d} \chi_{\mathbf{c}}(\hat{\mathbf{k}},l)\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},l)\frac{\partial \chi_{\mathbf{c}}(\hat{\mathbf{k}},l)}{\partial k_n}\right)\delta(k-l)\frac{d\mathbf{k}}{l^{d-1}}\\ &= \frac{1}{\Gamma_{d-1}}\int_{\mathbb{S}^{d-1}} (D(\hat{\mathbf{k}},l)\nabla\chi_{\mathbf{c}}(\hat{\mathbf{k}},l),\nabla\chi_{\mathbf{c}}(\hat{\mathbf{k}},l))_{\mathbb{R}^d}d\hat{\Omega}(\hat{\mathbf{k}}) \geq 0.\end{aligned}$$

The last equality holds after integration by parts because $D(\hat{\mathbf{k}},l)\hat{\mathbf{k}} = 0$. Moreover, the inequality appearing in the last line of (2.15) is strict. This can be seen as follows. Since the null-space of the matrix $D(\hat{\mathbf{k}},l)$ is one-dimensional and consists of the vectors parallel to $\hat{\mathbf{k}}$, in order for $(A\mathbf{c},\mathbf{c})_{\mathbb{R}^d}$ to vanish one needs that the gradient $\nabla\chi_{\mathbf{c}}(\hat{\mathbf{k}},l)$ is parallel to $\hat{\mathbf{k}}$ for all $\hat{\mathbf{k}} \in \mathbb{S}^{d-1}$. This, however, together with (2.14) would imply that $\hat{\mathbf{k}}\cdot\mathbf{c} = 0$ for all $\hat{\mathbf{k}}$, which is impossible.

The following theorem holds.

Theorem 2.5 *For every $0 < T_* < T < +\infty$ the re-scaled solution $\bar{\phi}_\gamma(t,\mathbf{x},\mathbf{k}) = \bar{\phi}(t/\gamma^2,\mathbf{x}/\gamma,\mathbf{k})$ of (2.10) converges as $\gamma \to 0$ in $C([T_*,T];L^\infty(\mathbb{R}^{2d}))$ to $w(t,\mathbf{x},\mathbf{k})$. Moreover, there exists a constant $C > 0$ so that we have*

$$\|w(t,\cdot) - \bar{\phi}_\gamma(t,\cdot)\|_{0,0} \leq C\left(\gamma T + \sqrt{\gamma}\right)\|\phi_0\|_{1,1} \qquad (2.16)$$

for all $T_* \leq t \leq T$.

Remark 2.6 In fact, as it will become apparent in the course of the proof, we have a stronger result, namely T_* can be made to vanish as $\gamma \to 0$. For instance, we can choose $T_* = \gamma^{3/2}$, see (4.16).

The proof of Theorem 2.5 is based on some classical asymptotic expansions and is quite straightforward. As an immediate corollary of Theorems 2.1 and 2.5 we obtain the following result, which is the main result of this paper.

Theorem 2.7 *Let ϕ_δ be solution of (2.8) with the initial data $\phi_\delta(0,\mathbf{x},\mathbf{k}) = \phi_0(\delta^{1+\alpha}\mathbf{x},\mathbf{k})$ and let $\bar{w}(t,\mathbf{x})$ be the solution of the diffusion equation (2.12) with the initial data $w(0,\mathbf{x},k) = \bar{\phi}_0(\mathbf{x},k)$. Then, there exists $\alpha_0 > 0$ and a constant $C > 0$ so that for all $0 \le \alpha < \alpha_0$ and all $0 < T_* \le T$ we have for all compact sets $K \subset \mathcal{A}(M)$:*

$$\sup_{(t,\mathbf{x},\mathbf{k})\in[T_*,T]\times K} \left| w(t,\mathbf{x},k) - \mathbb{E}\bar{\phi}_\delta(t,\mathbf{x},\mathbf{k}) \right| \le CT\delta^{\alpha_0-\alpha}, \tag{2.17}$$

where $\bar{\phi}_\delta(t,\mathbf{x},\mathbf{k}) := \phi_\delta\left(t/\delta^{1+2\alpha}, \mathbf{x}/\delta^{1+\alpha}, \mathbf{k}\right)$.

Theorem 2.7 shows that the movement of a particle in a weakly random quenched Hamiltonian is, indeed, approximated by a Brownian motion in the long time-large space limit, at least for times $T \ll \delta^{-\alpha_0}$. In fact, according to Remark 2.6 we can allow T_* to vanish as $\delta \to 0$ choosing $T_* = \delta^{3\alpha/2}$.

In the isotropic case when $R = R(|\mathbf{x}|,k)$ we may simplify the above expressions for the diffusion matrices D_{mn} and a_{mn}. In that case we have

$$\begin{aligned} D_{mn}(\hat{\mathbf{k}},k) &= -\frac{1}{2}\int_{-\infty}^{\infty} \frac{\partial^2 R(H_0'(k)s\hat{\mathbf{k}},k)}{\partial x_n \partial x_m}\, ds \\ &= -\int_0^\infty \left[\frac{k_n k_m}{k^2} R''(H_0'(k)s,k) + \left(\delta_{nm} - \frac{k_n k_m}{k^2}\right)\frac{R'(H_0'(k)s,k)}{H_0'(k)s}\right] ds \\ &= -\frac{1}{H_0'(k)}\int_0^\infty \frac{R'(s,k)}{s}\, ds \left(\delta_{nm} - \frac{k_n k_m}{k^2}\right), \end{aligned}$$

so that the matrix $[D_{mn}(\hat{\mathbf{k}},k)]$ has the form

$$D(\hat{\mathbf{k}},k) = D_0(k)\left(I - \hat{\mathbf{k}}\otimes\hat{\mathbf{k}}\right), \quad D_0(k) = -\frac{1}{H_0'(k)}\int_0^\infty \frac{R'(s,k)}{s} ds.$$

In that case the functions χ_j are given explicitly by

$$\chi_j(\hat{\mathbf{k}},k) = -\frac{|H_0'(k)|^2|k|^2\hat{k}_j}{(d-1)\bar{D}_0(k)}, \quad \bar{D}_0(k) = -\int_0^\infty \frac{R'(s,k)}{s} ds$$

and

$$a_{nm}(k) = \frac{|H_0'(k)|^3|k|^2}{\Gamma_{d-1}(d-1)\bar{D}_0(k)}\int_{\mathbb{S}^{d-1}} \hat{k}_n\hat{k}_m d\Omega(\hat{\mathbf{k}}) = \frac{|H_0'(k)|^3|k|^2}{d(d-1)\bar{D}_0(k)}\delta_{nm}.$$

2.6 A formal derivation of the momentum diffusion

We now recall how the diffusion operator in (2.10) can be derived in a quick formal way. We represent the solution of (2.8) as $\phi^\delta(t,\mathbf{x},\mathbf{k}) = \psi^\delta(\delta t,\delta\mathbf{x},\mathbf{k})$ and write an asymptotic multiple scale expansion for ψ^δ

$$\psi^\delta(t,\mathbf{x},\mathbf{k}) = \bar{\phi}(t,\mathbf{x},\mathbf{k}) + \sqrt{\delta}\phi_1\left(t,\mathbf{x},\frac{\mathbf{x}}{\delta},\mathbf{k}\right) + \delta\phi_2\left(t,\mathbf{x},\frac{\mathbf{x}}{\delta},\mathbf{k}\right) + \dots \tag{2.18}$$

We assume formally that the leading order term $\bar{\phi}$ is deterministic and independent of the fast variable $\mathbf{z} = \mathbf{x}/\delta$. We insert this expansion into (2.8) and obtain in the order $O\left(\delta^{-1/2}\right)$:

$$\nabla_{\mathbf{z}} H_1(\mathbf{z}, \mathbf{k}) \cdot \nabla_{\mathbf{k}} \bar{\phi} - H_0'(k) \hat{\mathbf{k}} \cdot \nabla_{\mathbf{z}} \phi_1 = 0. \tag{2.19}$$

Let $\theta \ll 1$ be a small positive regularization parameter that will be later sent to zero, and consider a regularized version of (2.19):

$$\frac{1}{H_0'(k)} \nabla_{\mathbf{z}} H_1(\mathbf{z}, \mathbf{k}) \cdot \nabla_{\mathbf{k}} \bar{\phi} - \hat{\mathbf{k}} \cdot \nabla_{\mathbf{z}} \phi_1 + \theta \phi_1 = 0,$$

Its solution is

$$\phi_1(\mathbf{z}, \mathbf{k}) = -\frac{1}{H_0'(k)} \int_0^\infty \sum_{m=1}^{d} \frac{\partial H_1(\mathbf{z} + s\hat{\mathbf{k}}, k)}{\partial z_m} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_m} e^{-\theta s} ds. \tag{2.20}$$

The next order equation becomes upon averaging

$$\frac{\partial \bar{\phi}}{\partial t} = \mathbb{E}\left(\frac{\partial H_1(\mathbf{z}, k)}{\partial k} \hat{\mathbf{k}} \cdot \nabla_{\mathbf{z}} \phi_1\right) - \mathbb{E}\left(\nabla_{\mathbf{z}} H_1(\mathbf{z}, k) \cdot \nabla_{\mathbf{k}} \phi_1\right) + H_0'(k) \hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} \bar{\phi}. \tag{2.21}$$

The first two terms on the right hand side above may be computed explicitly using expression (2.20) for ϕ_1:

$$\begin{aligned}
&\mathbb{E}\left(\frac{\partial H_1(\mathbf{z}, k)}{\partial k} \hat{\mathbf{k}} \cdot \nabla_{\mathbf{z}} \phi_1\right) - \mathbb{E}\left(\nabla_{\mathbf{z}} H_1(\mathbf{z}, k) \cdot \nabla_{\mathbf{k}} \phi_1\right) \\
&= -\mathbb{E}\left[\sum_{m,n=1}^{d} \frac{\partial H_1(\mathbf{z}, k)}{\partial k} \hat{k}_m \frac{\partial}{\partial z_m}\left(\frac{1}{H_0'(k)} \int_0^\infty \frac{\partial H_1(\mathbf{z} + s\hat{\mathbf{k}}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} e^{-\theta s} ds\right)\right] \\
&+\mathbb{E}\left[\sum_{m,n=1}^{d} \frac{\partial H_1(\mathbf{z}, k)}{\partial z_m} \frac{\partial}{\partial k_m}\left(\frac{1}{H_0'(k)} \int_0^\infty \frac{\partial H_1(\mathbf{z} + s\hat{\mathbf{k}}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} e^{-\theta s} ds\right)\right].
\end{aligned}$$

Using spatial stationarity of $H_1(\mathbf{z}, k)$ we may rewrite the above as

$$\begin{aligned}
&-\mathbb{E}\left[\sum_{m,n=1}^{d} \frac{\partial H_1(\mathbf{z}, k)}{\partial k} \hat{k}_m \frac{\partial}{\partial z_m}\left(\frac{1}{H_0'(k)} \int_0^\infty \frac{\partial H_1(\mathbf{z} + s\hat{\mathbf{k}}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} e^{-\theta s} ds\right)\right] \\
&-\mathbb{E}\left[\sum_{m,n=1}^{d} H_1(\mathbf{z}, k) \frac{\partial}{\partial z_m} \frac{\partial}{\partial k_m}\left(\frac{1}{H_0'(k)} \int_0^\infty \frac{\partial H_1(\mathbf{z} + s\hat{\mathbf{k}}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} e^{-\theta s} ds\right)\right] \\
&= -\sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left[\frac{1}{H_0'(k)} \int_0^\infty \mathbb{E}\left(H_1(\mathbf{z}, k) \frac{\partial^2 H_1(\mathbf{z} + s\hat{\mathbf{k}}, k)}{\partial z_n \partial z_m}\right) \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} e^{-\theta s} ds\right] \\
&= -\sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(\frac{1}{H_0'(k)} \int_0^\infty \frac{\partial^2 R(s\hat{\mathbf{k}}, k)}{\partial x_n \partial x_m} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} e^{-\theta s} ds\right) \\
&\to -\frac{1}{2} \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(\frac{1}{H_0'(k)} \int_{-\infty}^\infty \frac{\partial^2 R(s\hat{\mathbf{k}}, k)}{\partial x_n \partial x_m} \frac{\partial \bar{\phi}(t, \mathbf{x}, \mathbf{k})}{\partial k_n} ds\right), \text{ as } \theta \to 0^+.
\end{aligned}$$

We insert the above expression into (2.21) and obtain

$$\frac{\partial\bar{\phi}}{\partial t} = \sum_{m,n=1}^{d} \frac{\partial}{\partial k_n}\left(D_{nm}(\hat{\mathbf{k}},k)\frac{\partial\bar{\phi}}{\partial k_m}\right) + H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}\bar{\phi} \tag{2.22}$$

with the diffusion matrix $D(\hat{\mathbf{k}},k)$ as in (2.9). Observe that (2.22) is nothing but (2.10). However, the naive asymptotic expansion (2.18) may not be justified. The rigorous proof presented in the next section is based on a quite different method.

3 From the Liouville equation to the momentum diffusion. Estimation of the convergence rates: proof of Theorem 2.1

3.1 Outline of the proof

The basic idea of the proof of Theorem 2.1 is a modification of that of [1, 9]. We consider the trajectories corresponding to the Liouville equation (2.8) and introduce a stopping time, called τ_δ, that, among others, prevents near self-intersection of trajectories. This fact ensures that until the stopping time occurs the particle is "exploring a new territory" and, thanks to the strong mixing properties of the medium, "memory effects" are lost. Therefore, roughly speaking, until the stopping time the process is approximately characterized by the Markov property. Furthermore, since the amplitude of the random Hamiltonian is not strong enough to destroy the continuity of its path, it becomes a diffusion in the limit, as $\delta\to 0$. We introduce also an augmented process that follows the trajectories of the Hamiltonian flow until the stopping time τ_δ and becomes a diffusion after $t=\tau_\delta$. We show that the law of the augmented process is close to the law of a diffusion, see Proposition 3.4, with an explicit error bound. We also prove that the stopping time tends to infinity as $\delta\to 0$, once again with the error bound that is proved in Theorem 3.6. The combination of these two results allows us to estimate the difference between the solutions of the Liouville and the diffusion equations in a rather straightforward manner (see Section 3.9): they are close until the stopping time as the law of the diffusion is always close to that of the augmented process, while the latter coincides with the true process until τ_δ. On the other hand, the fact that $\tau_\delta\to\infty$ as $\delta\to 0$ shows that with a large probability the augmented process is close to the true process. This combination finishes the proof.

3.2 The random characteristics corresponding to (2.8)

Consider the motion of a particle governed by a Hamiltonian system of equations

$$\begin{cases} \frac{d\mathbf{z}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{dt} = (\nabla_{\mathbf{k}}H_\delta)\left(\frac{\mathbf{z}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{\delta},\mathbf{m}^{(\delta)}(t;\mathbf{x},\mathbf{k})\right) \\ \frac{d\mathbf{m}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{dt} = -\frac{1}{\sqrt{\delta}}(\nabla_{\mathbf{z}}H_\delta)\left(\frac{\mathbf{z}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{\delta},\mathbf{m}^{(\delta)}(t;\mathbf{x},\mathbf{k})\right) \\ \mathbf{z}^{(\delta)}(0;\mathbf{x},\mathbf{k})=\mathbf{x}, \quad \mathbf{m}^{(\delta)}(0;\mathbf{x},\mathbf{k})=\mathbf{k}, \end{cases} \tag{3.1}$$

where the Hamiltonian $H_\delta(\mathbf{x},\mathbf{k}) := H_0(k)+\sqrt{\delta}H_1(\mathbf{x},k)$, $k=|\mathbf{k}|$. The trajectories of (3.1) are the characteristics of the Liouville equation (2.8). The hypotheses made in Section 2 imply that the trajectory $(\mathbf{z}^{(\delta)}(t;\mathbf{x},\mathbf{k}),\mathbf{m}^{(\delta)}(t;\mathbf{x},\mathbf{k}))$ necessarily lies in $\mathcal{C}(T,\delta)$ for each $T>0$, $\delta\in(0,\delta_*(M)]$, provided that the initial data $(\mathbf{x},\mathbf{k})\in\mathcal{A}(M)$. Indeed, it follows from the Hamiltonian structure of (3.1) that the Hamiltonian $H_\delta(\mathbf{x},m)=H_0(m)+\sqrt{\delta}H_1(\mathbf{z},m)$ must be conserved along the trajectory.

Hence, the definition (2.7) implies that $M_\delta^{-1} \le |\mathbf{m}^{(\delta)}(\cdot;\mathbf{x},\mathbf{k})| \le M_\delta$. We denote by $Q^\delta_{s,\mathbf{x},\mathbf{k}}(\cdot)$ the law over $\mathcal{C}$ of the process corresponding to (3.1) starting at $t = s$ from $(\mathbf{x},\mathbf{k})$ (this law is actually supported in $\mathcal{C}(\delta)$). We shall omit writing the subscript s when it equals to 0.

3.3 The stopping times

We now define the stopping time τ_δ, described in Section 3.1, that prevents the trajectories of (3.1) to have near self-intersections (recall that the intent of the stopping time is to prevent any "memory effects" of the trajectories). As we have already mentioned, we will later show that the probability of the event $[\,\tau_\delta < T\,]$ for a fixed $T > 0$ goes to zero, as $\delta \to 0$.

Let $0 < \epsilon_1 < \epsilon_2 < 1/2$, $\epsilon_3 \in (0, 1/2 - \epsilon_2)$, $\epsilon_4 \in (1/2, 1 - \epsilon_1 - \epsilon_2)$ be small positive constants that will be further determined later and set

$$N = [\delta^{-\epsilon_1}], \quad p = [\delta^{-\epsilon_2}], \quad q = p\,[\delta^{-\epsilon_3}], \quad N_1 = Np\,[\delta^{-\epsilon_4}]. \tag{3.2}$$

We will specify additional restrictions on the constants ϵ_j as the need for such constraints arises. However, the basic requirement is that ϵ_i, $i = 1,2,3$ should be sufficiently small and ϵ_4 is bigger than 1/2, less than one and can be made as close to one as we would need it. It is important that $\epsilon_1 < \epsilon_2$ so that $N \ll p$ when $\delta \ll 1$. We introduce the following $(\mathcal{M}^t)_{t\ge 0}$–stopping times. Let $t_k^{(p)} := kp^{-1}$ be a mesh of times, and $\pi \in \mathcal{C}$ be a path. We define the "violent turn" stopping time

$$\begin{aligned} S_\delta(\pi) := \inf \Bigg[& t \ge 0 : \text{ for some } k \ge 0 \text{ we have } t \in \left[t_k^{(p)}, t_{k+1}^{(p)}\right) \text{ and} \\ & \hat{K}(t_{k-1}^{(p)}) \cdot \hat{K}(t) \le 1 - \frac{1}{N}, \text{ or } \hat{K}\left(t_k^{(p)} - \frac{1}{N_1}\right) \cdot \hat{K}(t) \le 1 - \frac{1}{N} \Bigg], \end{aligned} \tag{3.3}$$

where by convention we set $\hat{K}(-1/p) := \hat{K}(0)$. Note that with the above choice of ϵ_4 we have $\hat{K}\left(t_k^{(p)} - 1/N_1\right) \cdot \hat{K}(t_k^{(p)}) > 1 - 1/N$, provided that $\delta \in (0, \delta_0]$ and δ_0 is sufficiently small. We adopt in (3.3) a customary convention that the infimum of an empty set equals $+\infty$. The stopping time S_δ is triggered when the trajectory performs a sudden turn – this is undesirable as the trajectory may then return back to the region it has already visited and create correlations with the past.

For each $t \ge 0$, we denote by $\mathfrak{X}_t(\pi) := \bigcup\limits_{0\le s\le t} X(s;\pi)$ the trace of the spatial component of the path π up to time t, and by $\mathfrak{X}_t(q;\pi) := [\mathbf{x} : \operatorname{dist}(\mathbf{x}, \mathfrak{X}_t(\pi)) \le 1/q]$ a tubular region around the path. We introduce the stopping time

$$U_\delta(\pi) := \inf \left[t \ge 0 : \exists\, k \ge 1 \text{ and } t \in [t_k^{(p)}, t_{k+1}^{(p)}) \text{ for which } X(t) \in \mathfrak{X}_{t_{k-1}^{(p)}}(q) \right]. \tag{3.4}$$

It is associated with the return of the X component of the trajectory to the tube around its past – this is again an undesirable way to create correlations with the past. Finally, we set the stopping time

$$\tau_\delta(\pi) := S_\delta(\pi) \wedge U_\delta(\pi). \tag{3.5}$$

3.4 The cut-off functions and the corresponding dynamics

Let $M > 0$ be fixed and p, q, N, N_1 be the positive integers defined in Section 3.3. We define now several auxiliary functions that will be used to introduce the cut-offs in the dynamics. These cut-offs will ensure that the particle moving under the modified dynamics will avoid self-intersections, will have no violent turns and the changes of its momentum will be under control. In addition, up to

the stopping time τ_δ the motion of the particle will coincide with the motion under the original Hamiltonian flow.

Let $a_1 = 2$ and $a_2 = 3/2$. The functions $\psi_j : \mathbb{R}^d \times \mathbb{S}_1^{d-1} \to [0,1]$, $j = 1,2$ are of C^∞ class and satisfy

$$\psi_j(\mathbf{k},\mathbf{l}) = \begin{cases} 1, & \text{if} \quad \hat{\mathbf{k}}\cdot\mathbf{l} \geq 1 - 1/N \qquad \text{and} \qquad M_\delta^{-1} \leq |\mathbf{k}| \leq M_\delta \\ \\ 0, & \text{if} \quad \hat{\mathbf{k}}\cdot\mathbf{l} \leq 1 - a_j/N, \quad \text{or} \quad |\mathbf{k}| \leq (2M_\delta)^{-1}, \quad \text{or} \quad |\mathbf{k}| \geq 2M_\delta. \end{cases} \tag{3.6}$$

One can construct ψ_j in such a way that for arbitrary nonnegative integers m, n it is possible to find a constant $C_{m,n}$ for which $\|\psi_j\|_{m,n} \leq C_{m,n} N^{m+n}$. The cut-off function

$$\Psi(t,\mathbf{k};\pi) := \begin{cases} \psi_1\left(\mathbf{k}, \hat{K}\left(t_{k-1}^{(p)}\right)\right)\psi_2\left(\mathbf{k}, \hat{K}\left(t_k^{(p)} - 1/N_1\right)\right) & \text{for } t \in [t_k^{(p)}, t_{k+1}^{(p)}) \text{ and } k \geq 1 \\ \psi_2(\mathbf{k}, \hat{K}(0)) & \text{for } t \in [0, t_1^{(p)}) \end{cases} \tag{3.7}$$

will allow us to control the direction of the particle motion over each interval of the partition as well as not to allow the trajectory to escape to the regions where the change of the size of the velocity can be uncontrollable.

Let $\phi : \mathbb{R}^d \times \mathbb{R}^d \to [0,1]$ be a function of the C^∞ class that satisfies $\phi(\mathbf{y},\mathbf{x}) = 1$, when $|\mathbf{y}-\mathbf{x}| \geq 3/q$ and $\phi(\mathbf{y},\mathbf{x}) = 0$, when $|\mathbf{y}-\mathbf{x}| \leq 2/q$. Again, in this case we can construct ϕ in such a way that $\|\phi\|_{m,n} \leq Cq^{m+n}$ for arbitrary integers m, n and a suitably chosen constant C. The function $\phi_k : \mathbb{R}^d \times \mathcal{C} \to [0,1]$ for a fixed path π is given by

$$\phi_k(\mathbf{y};\pi) = \prod_{0 \leq l/q \leq t_{k-1}^{(p)}} \phi\left(\mathbf{y}, X\left(\frac{l}{q}\right)\right). \tag{3.8}$$

We set

$$\Phi(t,\mathbf{y};\pi) := \begin{cases} 1, & \text{if } 0 \leq t < t_1^{(p)} \\ \phi_k(\mathbf{y};\pi), & \text{if } t_k^{(p)} \leq t < t_{k+1}^{(p)}. \end{cases} \tag{3.9}$$

The function Φ shall be used to modify the dynamics of the particle in order to avoid a possibility of near self-intersections of its trajectory.

For a given $t \geq 0$, $(\mathbf{y},\mathbf{k}) \in \mathbb{R}_*^{2d}$ and $\pi \in \mathcal{C}$ let us denote $\Theta(t,\mathbf{y},\mathbf{k};\pi) := \Psi(t,\mathbf{k};\pi)\Phi(t,\mathbf{y};\pi)$. The following lemma can be verified by a direct calculation.

Lemma 3.1 *Let (β_1,β_2) be a multi-index with nonnegative integer valued components, $m = |\beta_1| + |\beta_2|$. There exists a constant C depending only on m and M such that $|\partial_{\mathbf{y}}^{\beta_1}\partial_{\mathbf{k}}^{\beta_2}\Theta(t,\mathbf{y},\mathbf{k};\pi)| \leq CT^{|\beta_1|}q^{2|\beta_1|}N^{|\beta_2|}$ for all $t \in [0,T]$, $(\mathbf{y},\mathbf{k}) \in \mathcal{A}(2M)$, $\pi \in \mathcal{C}$.*

Finally, let us set

$$F_\delta(t,\mathbf{y},\mathbf{l};\pi,\omega) = \Theta(t,\delta\mathbf{y},\mathbf{l};\pi)\nabla_{\mathbf{y}} H_1(\mathbf{y},|\mathbf{l}|;\omega). \tag{3.10}$$

For a fixed $(\mathbf{x},\mathbf{k}) \in \mathbb{R}_*^{2d}$, $\delta > 0$ and $\omega \in \Omega$ we consider the modified particle dynamics with the cut-off that is described by the stochastic process $(\boldsymbol{y}^{(\delta)}(t;\mathbf{x},\mathbf{k},\omega), \boldsymbol{l}^{(\delta)}(t;\mathbf{x},\mathbf{k},\omega))_{t\geq 0}$ whose paths are the solutions of the following equation

$$\begin{cases} \frac{d\boldsymbol{y}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{dt} = \left[H_0'(|\boldsymbol{l}^{(\delta)}(t;\mathbf{x},\mathbf{k})|) + \sqrt{\delta}\,\partial_l H_1\left(\frac{\boldsymbol{y}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{\delta}, |\boldsymbol{l}^{(\delta)}(t;\mathbf{x},\mathbf{k})|\right)\right]\hat{\boldsymbol{l}}^{(\delta)}(t;\mathbf{x},\mathbf{k},) \\ \\ \frac{d\boldsymbol{l}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{dt} = -\frac{1}{\sqrt{\delta}}F_\delta\left(t, \frac{\boldsymbol{y}^{(\delta)}(t;\mathbf{x},\mathbf{k})}{\delta}, \boldsymbol{l}^{(\delta)}(t;\mathbf{x},\mathbf{k}); \boldsymbol{y}^{(\delta)}(\cdot;\mathbf{x},\mathbf{k}), \boldsymbol{l}^{(\delta)}(\cdot;\mathbf{x},\mathbf{k})\right) \\ \\ \boldsymbol{y}^{(\delta)}(0;\mathbf{x},\mathbf{k}) = \mathbf{x}, \quad \boldsymbol{l}^{(\delta)}(0;\mathbf{x},\mathbf{k}) = \mathbf{k}. \end{cases} \tag{3.11}$$

We will denote by $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}$ the law of the modified process $(\boldsymbol{y}^{(\delta)}(\cdot;\mathbf{x},\mathbf{k}),\boldsymbol{l}^{(\delta)}(\cdot;\mathbf{x},\mathbf{k}))$ over $\mathcal{C}$ for a given $\delta>0$ and by $\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}$ the corresponding expectation. We assume that the initial momentum $\mathbf{k}\in A(M)$. From the construction of the cut-offs we immediately conclude that

$$\hat{\boldsymbol{l}}^{(\delta)}(t)\cdot\hat{\boldsymbol{l}}^{(\delta)}(t^{(p)}_{k-1})\geq 1-\frac{2}{N},\quad t\in[t^{(p)}_{k-1},t^{(p)}_{k+1}),\quad \forall\, k\geq 0. \tag{3.12}$$

3.5 Some consequences of the mixing assumption

For any $t\geq 0$ we denote by $\mathcal{F}_t$ the σ-algebra generated by $(\boldsymbol{y}^{(\delta)}(s),\boldsymbol{l}^{(\delta)}(s))$, $s\leq t$. Here we suppress, for the sake of abbreviation, writing the initial data in the notation of the trajectory. In this section we assume that $M>0$ is fixed, $X_1,X_2:(\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^{d^2})^2\to\mathbb{R}$ are certain continuous functions, Z is a random variable and g_1,g_2 are $\mathbb{R}^d\times[M^{-1},M]$-valued random vectors. We suppose further that Z,g_1,g_2, are $\mathcal{F}_t$-measurable, while $\tilde{X}_1,\tilde{X}_2$ are random fields of the form

$$\tilde{X}_i(\mathbf{x},k)=X_i\left(\left(\partial_k^j H_1(\mathbf{x},k),\nabla_{\mathbf{x}}\partial_k^j H_1(\mathbf{x},k),\nabla^2_{\mathbf{x}}\partial_k^j H_1(\mathbf{x},k)\right)_{j=0,1}\right).$$

For $i=1,2$ we denote $g_i:=(g_i^{(1)},g_i^{(2)})$ where $g_i^{(1)}\in\mathbb{R}^d$ and $g_i^{(2)}\in[M^{-1},M]$. We also let

$$U(\theta_1,\theta_2):=\mathbb{E}\left[\tilde{X}_1(\theta_1)\tilde{X}_2(\theta_2)\right],\quad \theta_1,\theta_2\in\mathbb{R}^d\times[M^{-1},M]. \tag{3.13}$$

The following mixing lemma is useful in formalizing the "memory loss effect" and can be proved in the same way as Lemmas 5.2 and 5.3 of [1].

Lemma 3.2 *(i) Assume that $r,t\geq 0$ and*

$$\inf_{u\leq t}\left|g_i^{(1)}-\frac{\boldsymbol{y}^{(\delta)}(u)}{\delta}\right|\geq\frac{r}{\delta}, \tag{3.14}$$

$\mathbb{P}$–a.s. on the set $Z\neq 0$ for $i=1,2$. Then, we have

$$\left|\mathbb{E}\left[\tilde{X}_1(g_1)\tilde{X}_2(g_2)Z\right]-\mathbb{E}\left[U(g_1,g_2)Z\right]\right|\leq 2\phi\left(\frac{r}{2\delta}\right)\|X_1\|_{L^\infty}\|X_2\|_{L^\infty}\|Z\|_{L^1(\Omega)}. \tag{3.15}$$

(ii) Let $\mathbb{E}X_1(\mathbf{0},k)=0$ for all $k\in[M^{-1},M]$. Furthermore, we assume that g_2 satisfies (3.14),

$$\inf_{u\leq t}\left|g_1^{(1)}-\frac{\boldsymbol{y}^{(\delta)}(u)}{\delta}\right|\geq\frac{r+r_1}{\delta} \tag{3.16}$$

and $|g_1^{(1)}-g_2^{(1)}|\geq r_1\delta^{-1}$ for some $r_1\geq 0$, $\mathbb{P}$-a.s. on the event $Z\neq 0$. Then, we have

$$\left|\mathbb{E}\left[\tilde{X}_1(g_1)\tilde{X}_2(g_2)\,Z\right]-\mathbb{E}\left[U(g_1,g_2)Z\right]\right|\leq C\phi^{1/2}\left(\frac{r}{2\delta}\right)\phi^{1/2}\left(\frac{r_1}{2\delta}\right)\|X_1\|_{L^\infty}\|X_2\|_{L^\infty}\|Z\|_{L^1(\Omega)} \tag{3.17}$$

for some absolute constant $C>0$. Here the function U is given by (3.13).

3.6 The momentum diffusion

Let $\mathbf{k}(t)$ be a diffusion, starting at $\mathbf{k} \in \mathbb{R}^d_*$ at $t = 0$, with the generator of the form

$$\begin{aligned}\mathcal{L}F(\mathbf{k}) &= \sum_{m,n=1}^{d} D_{mn}(\hat{\mathbf{k}}, |\mathbf{k}|)\partial^2_{k_m,k_n}F(\mathbf{k}) + \sum_{m=1}^{d} E_m(\hat{\mathbf{k}}, |\mathbf{k}|)\partial_{k_m}F(\mathbf{k}) \\ &= \sum_{m,n=1}^{d} \partial_{k_m}\left(D_{m,n}(\hat{\mathbf{k}}, |\mathbf{k}|)\partial_{k_n}F(\mathbf{k})\right), \quad F \in C_0^\infty(\mathbb{R}^d_*).\end{aligned} \tag{3.18}$$

Here the diffusion matrix is given by (2.9) and the drift vector is

$$E_m(\hat{\mathbf{k}}, l) = -\frac{1}{H_0'(l)l}\sum_{n=1}^{d}\int_0^{+\infty} s\frac{\partial^3 R(s\hat{\mathbf{k}}, l)}{\partial x_m \partial x_n^2}\, ds, \quad m = 1, \ldots, d.$$

Employing exactly the same argument as the one used in Section 4 of [1] it can be easily seen that this diffusion is supported on $\mathbb{S}^{d-1}_k$, where $k = |\mathbf{k}|$. Moreover, it is non-degenerate on the sphere, for instance, under the assumption (2.6), cf. Proposition 4.3 of ibid.

Let $\mathfrak{Q}_{\mathbf{x},\mathbf{k}}$ be the law of the process $(\mathbf{x}(t), \mathbf{k}(t))$ that starts at $t = 0$ from $(\mathbf{x}, \mathbf{k})$ given by $\mathbf{x}(t) = \mathbf{x} + \int_0^t H_0'(|\mathbf{k}(s)|)\hat{\mathbf{k}}(s)ds$, where $\mathbf{k}(t)$ is the diffusion described by (3.18). This process is a degenerate diffusion whose generator is given by

$$\tilde{\mathcal{L}}F(\mathbf{x}, \mathbf{k}) = \mathcal{L}_{\mathbf{k}}F(\mathbf{x}, \mathbf{k}) + H_0'(|\mathbf{k}|)\,\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}}F(\mathbf{x}, \mathbf{k}), \quad F \in C_0^\infty(\mathbb{R}^{2d}_*). \tag{3.19}$$

Here the notation $\mathcal{L}_{\mathbf{k}}$ stresses that the operator $\mathcal{L}$ defined in (3.18) acts on the respective function in the $\mathbf{k}$ variable. We denote by $\mathfrak{M}_{\mathbf{x},\mathbf{k}}$ the expectation corresponding to the path measure $\mathfrak{Q}_{\mathbf{x},\mathbf{k}}$.

3.7 The augmented process

The following construction of the augmentation of path measures has been carried out in Section 6.1 of [14]. Let $s \geq 0$ be fixed and $\pi \in \mathcal{C}$. Then, according to Lemma 6.1.1 of ibid. there exists a unique probability measure, that is denoted by $\delta_\pi \otimes_s \mathfrak{Q}_{X(s),K(s)}$, such that for any pair of events $A \in \mathcal{M}^s$, $B \in \mathcal{M}$ we have $\delta_\pi \otimes_s \mathfrak{Q}_{X(s),K(s)}[A] = \mathbf{1}_A(\pi)$ and $\delta_\pi \otimes_s \mathfrak{Q}_{X(s),K(s)}[\theta_s(B)] = \mathfrak{Q}_{X(s),K(s)}[B]$. The following result is a direct consequence of Theorem 6.2.1 of [14].

Proposition 3.3 *There exists a unique probability measure $R^{(\delta)}_{\mathbf{x},\mathbf{k}}$ on $\mathcal{C}$ such that $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A] := Q^{(\delta)}_{\mathbf{x},\mathbf{k}}[A]$ for all $A \in \mathcal{M}^{\tau_\delta}$ and the regular conditional probability distribution of $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[\,\cdot\,|\mathcal{M}^{\tau_\delta}]$ is given by $\delta_\pi \otimes_{\tau_\delta(\pi)} \mathfrak{Q}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}$, $\pi \in \mathcal{C}$. This measure shall be also denoted by $Q^{(\delta)}_{\mathbf{x},\mathbf{k}} \otimes_{\tau_\delta} \mathfrak{Q}_{X(\tau_\delta),K(\tau_\delta)}$.*

Note that for any $(\mathbf{x}, \mathbf{k}) \in \mathcal{A}(M)$ and $A \in \mathcal{M}^{\tau_\delta}$ we have

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A] = Q^{(\delta)}_{\mathbf{x},\mathbf{k}}[A] = \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[A], \tag{3.20}$$

that is, the law of the augmented process coincides with that of the true process, and of the modified process with the cut-offs until the stopping time τ_δ. Hence, according to the uniqueness part of Proposition 3.3, in such a case $Q^{(\delta)}_{\mathbf{x},\mathbf{k}} \otimes_{\tau_\delta} \mathfrak{Q}_{X(\tau_\delta),K(\tau_\delta)} = \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}} \otimes_{\tau_\delta} \mathfrak{Q}_{X(\tau_\delta),K(\tau_\delta)}$. We denote by $E^{(\delta)}_{\mathbf{x},\mathbf{k}}$ the expectation with respect to the augmented measure described by the above proposition. Let also $R^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}$, $E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}$ denote the respective conditional law and expectation obtained by conditioning $R^{(\delta)}_{\mathbf{x},\mathbf{k}}$ on $\mathcal{M}^{\tau_\delta}$.

The following proposition is of crucial importance for us, as it shows that the law of the augmented process is close to that of the momentum diffusion as $\delta \to 0$. To abbreviate the notation we let

$$N_t(G) := G(t, X(t), K(t)) - G(0, X(0), K(0)) - \int_0^t (\partial_\varrho + \tilde{\mathcal{L}}) G(\varrho, X(\varrho), K(\varrho)))\, d\varrho$$

for any $G \in C_b^{1,1,3}([0, +\infty) \times \mathbb{R}^{2d}_*)$ and $t \geq 0$.

Proposition 3.4 *Suppose that* $(\mathbf{x}, \mathbf{k}) \in \mathcal{A}(M)$ *and* $\zeta \in C_b((\mathbb{R}^{2d}_*)^n)$ *is nonnegative. Let* $\gamma_0 \in (0, 1/2)$ *and let* $0 \leq t_1 < \cdots < t_n \leq T_* \leq t < v \leq T$. *We assume further that* $v - t \geq \delta^{\gamma_0}$. *Then, there exist constants* γ_1, C *such that for any function* $G \in C^{1,1,3}([T_*, T] \times \mathbb{R}^{2d}_*)$ *we have*

$$\left| E^{(\delta)}_{\mathbf{x},\mathbf{k}} \left\{ [N_v(G) - N_t(G)]\, \tilde{\zeta} \right\} \right| \leq C \delta^{\gamma_1} (v - t) \|G\|^{[T_*, T]}_{1,1,3} T^2 E^{(\delta)}_{\mathbf{x},\mathbf{k}} \tilde{\zeta}. \tag{3.21}$$

Here $\tilde{\zeta}(\pi) := \zeta(X(t_1), K(t_1), \ldots, X(t_n), K(t_n))$, $\pi \in \mathcal{C}(T, \delta)$. *The choice of the constants* γ_1, C *does not depend on* $(\mathbf{x}, \mathbf{k})$, $\delta \in (0, 1]$, ζ, *times* $t_1, \ldots, t_n, T_*, T, v, t$, *or the function* G.

Proof. Let $0 = s_0 \leq s_1 \leq \ldots \leq s_n \leq t$ and $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}^{2d}_*)$ be Borel sets. We denote $A_0 := \mathcal{C}$ and for any $k \in \{1, \ldots, n\}$, $s \leq s_k$ we define the events

$$A_k := [\pi :\ (X(s_1), K(s_1)) \in B_1, \ldots, (X(s_k), K(s_k)) \in B_k]$$

and their shifted counterparts

$$A^{(s)}_k := [\pi :\ (X(s_k - s), K(s_k - s)) \in B_k, \ldots, (X(s_n - s), K(s_n - s)) \in B_n].$$

For $(\mathbf{x}, \mathbf{k}) \in \mathbb{R}^{2d}_*$, $\pi \in \mathcal{C}$ and $G \in C^{1,1,2}([0, +\infty) \times \mathbb{R}^{2d}_*)$ we let

$$\begin{aligned}\widehat{\mathcal{L}}_t G(t, \mathbf{x}, \mathbf{k}; \pi) := {} & H_0'(|\mathbf{k}|)\, \hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} G(t, \mathbf{x}, \mathbf{k}) + \Theta^2(t, X(t), K(t); \pi) \mathcal{L}_{\mathbf{k}} G(t, \mathbf{x}, \mathbf{k}) \\ & - \Theta(t, X(t), K(t); \pi) \sum_{m,n=1}^{d} \partial_{K_m} \Theta(t, X(t), K(t); \pi) D_{m,n}(\hat{\mathbf{k}}, |\mathbf{k}|) \partial_{k_n} G(t, \mathbf{x}, \mathbf{k})\end{aligned}$$

and

$$\widehat{N}_t(G) := G(t, X(t), K(t)) - G(0, X(0), K(0)) - \int_0^t (\partial_\varrho + \widehat{\mathcal{L}}_\varrho) G(\varrho, X(\varrho), K(\varrho); \pi)\, d\varrho.$$

It follows from the definition of the stopping time $\tau_\delta(\pi)$ and the cut-off function Θ that

$$\nabla_K \Theta(t, X(t), K(t); \pi) = \mathbf{0}, \quad t \in [0, \tau_\delta(\pi)],$$

hence

$$\widehat{\mathcal{L}}_t G(t, X(t), K(t); \pi) = \tilde{\mathcal{L}} G(t, X(t), K(t); \pi), \quad t \in [0, \tau_\delta(\pi)].$$

We need the following result.

Lemma 3.5 *Suppose that* $(\mathbf{x}, \mathbf{k}) \in \mathcal{A}(M)$ *and* $\zeta \in C_b((\mathbb{R}^{2d}_*)^n)$ *is nonnegative. Let* $\gamma_0' \in (0, 1)$, $0 \leq t_1 < \cdots < t_n \leq T_* \leq t < v \leq T$ *and* $t - T_* \geq \delta^{\gamma_0'}$. *Then, there exist constants* γ_1', $C' > 0$ *such that for any function* $G \in C^{1,1,3}([T_*, T] \times \mathbb{R}^{2d}_*)$ *we have*

$$\left| \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}} \left\{ [\widehat{N}_v(G) - \widehat{N}_t(G)] \tilde{\zeta} \right\} \right| \leq C' \delta^{\gamma_1'} (v - t) \|G\|^{[T_*, T]}_{1,1,3} T^2 \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}} \tilde{\zeta}. \tag{3.22}$$

The choice of the constants γ_1', C' *does not depend on* $(\mathbf{x}, \mathbf{k})$, $\delta \in (0, 1]$, *times* $t_1, \ldots, t_n, T_*$, T, v, t, *or function* G.

The proof of this lemma follows very closely the argument presented in Section 5.3 of [1] and we postpone it until the Appendix. In the meantime we apply this result to conclude the proof of Proposition 3.4. We write

$$
\begin{aligned}
E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[N_v(G)-N_{v\wedge\tau_\delta(\pi)}(G),A_n] = &\sum_{p=0}^{n-1}\mathbf{1}_{[s_p,s_{p+1})}(\tau_\delta(\pi))\mathbf{1}_{A_p}(\pi)\mathfrak{M}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G),A_{p+1}^{(\tau_\delta(\pi))}] \\
&+\mathbf{1}_{[s_n,v)}(\tau_\delta(\pi))\mathbf{1}_{A_n}(\pi)\mathfrak{M}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G)]. \qquad (3.23)
\end{aligned}
$$

When $\tau_\delta(\pi)\in[s_p,s_{p+1})$ we obviously have

$$
\mathfrak{M}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G),A_{p+1}^{(\tau_\delta(\pi))}]=\mathfrak{M}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[N_{t-\tau_\delta(\pi)}(G),A_{p+1}^{(\tau_\delta(\pi))}]
$$

and $\mathfrak{M}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G)]=0$. Hence the left hand side of (3.23) equals

$$
\begin{aligned}
&\sum_{p=0}^{n-1}\mathbf{1}_{[s_p,s_{p+1})}(\tau_\delta(\pi))\mathbf{1}_{A_p}(\pi)\mathfrak{M}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[N_{t-\tau_\delta(\pi)}(G),A_{p+1}^{(\tau_\delta(\pi))}] \qquad (3.24)\\
&=E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[N_t(G)-N_{t\wedge\tau_\delta(\pi)}(G),A_n].
\end{aligned}
$$

We conclude from (3.23), (3.24) that

$$
\begin{aligned}
E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[N_v(G),A_n] &= E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[N_{v\wedge\tau_\delta(\pi)}(G)+N_t(G)-N_{t\wedge\tau_\delta(\pi)}(G),A_n] \qquad (3.25)\\
&=E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[N_{(v\wedge\tau_\delta(\pi))\vee t}(G),A_n]
\end{aligned}
$$

and therefore

$$
\begin{aligned}
&E^{(\delta)}_{\mathbf{x},\mathbf{k}}[N_v(G),A_n]=E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[N_{(v\wedge\tau_\delta(\pi))\vee t}(G),A_n]\right] \qquad (3.26)\\
&=E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}\left[N_{(v\wedge\tau_\delta(\pi))\vee t}(G),A_n\right],\tau_\delta(\pi)\le t\right]+E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[E^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}\left[N_{(v\wedge\tau_\delta(\pi))\vee t}(G),A_n\right],\,\tau_\delta(\pi)>t\right].
\end{aligned}
$$

The first term on the utmost right hand side of (3.26) equals $E^{(\delta)}_{\mathbf{x},\mathbf{k}}[N_t(G),A_n,\,\tau_\delta\le t]$, while the second one equals $\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{(v\wedge\tau_\delta)\vee t}(G),B\right]$. Here $B:=A_n\cap[\tau_\delta>t]$ is an $\mathcal{M}_t$–measurable event. Suppose that $\gamma_0'\in(\gamma_0+1/2,1)$ and let $L:=[\delta^{-\gamma_0'}]$ be yet another mesh size parameter. We define

$$
\sigma:=L^{-1}[([L(v\wedge\tau_\delta)]+2)\vee([Lt]+2)]
$$

and note that

$$
\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}[N_\sigma(G),\,B]=\sum_{p=[Lt]+2}^{[Lv]+2}\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{p/L}(G),\,B,\,\sigma=\frac{p}{L}\right] \qquad (3.27)
$$

Representing the event $[\sigma=p/L]$ as the difference of $[\sigma\ge p/L]$ and $[\sigma\ge(p+1)/L]$ (note that $[\sigma\ge([Lv]+3)/L]=\emptyset$) and grouping the terms of the sum that correspond to the same index p we obtain that the right hand side of (3.27) equals

$$
\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{([Lt]+2)/L}(G),\,B\,\right]+\sum_{p=[Lt]+2}^{[Lv]+2}\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{p+1/L}(G)-N_{p/L}(G),\,B,\,\sigma\ge\frac{p+1}{L}\right]. \qquad (3.28)
$$

Since the event $B \cap [\sigma \geq (p+1)/L]$ is $\mathcal{M}^{(p-1)/L}$-measurable, from Lemma 3.5 we conclude that the absolute value of each term appearing under the summation sign in (3.28) can be estimated by $C'\|G\|_{1,1,3}\delta^{\gamma_1'}L^{-1}\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[B]$ which implies

$$\left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_\sigma(G),\, B\right] - \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{([Lt]+2)L^{-1}}(G),\, B\right]\right| \leq C'\delta^{\gamma_1'}\|G\|^{[T_*,T]}_{1,1,3}T^2\,\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[B]\,\frac{[Lv]+1-[Lt]}{L}.$$

A direct calculation using formulas (3.1) allows us to conclude also that both $|N_\sigma(G) - N_{(v\wedge\tau_\delta)\vee t}(G)|$ and $|N_{([Lt]+2)L^{-1}}(G) - N_t(G)|$ are estimated by $C\|G\|^{[T_*,T]}_{1,1,3}\delta^{\gamma_0'-1/2}$. Hence, (since $\gamma_0' > 1/2 + \gamma_0$)

$$\begin{aligned}
&\left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{(v\wedge\tau_\delta)\vee t}(G),\, B\right] - \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_t(G),\, B\right]\right| \leq \left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_\sigma(G) - N_{(v\wedge\tau_\delta)\vee t}(G),\, B\right]\right| \\
&+ \left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_\sigma(G),\, B\right] - \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{([Lt]+2)L^{-1}}(G),\, B\right]\right| + \left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[N_{([Lt]+2)L^{-1}}(G) - N_t,\, B\right]\right| \\
&\leq C\delta^{\gamma_1}\|G\|^{[T_*,T]}_{1,1,3}T^2\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[B]\,(v-t)\vee\delta^{\gamma_0} \qquad (3.29)
\end{aligned}$$

for a certain constant $C > 0$ and $\gamma_1 := \min[\gamma_0' - \gamma_0 - 1/2, \gamma_1']$. From (3.26), (3.29) and the observation just below (3.26), we obtain

$$\left|E^{(\delta)}_{\mathbf{x},\mathbf{k}}[N_v(G) - N_t(G), A_n]\right| \leq C\delta^{\gamma_1}\|G\|^{[T_*,T]}_{1,1,3}T^2R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A_n]\,(v-t)\vee\delta^{\gamma_0}$$

for a certain constant $C > 0$ and the conclusion of Proposition 3.4 follows. □

3.8 An estimate of the stopping time

The purpose of this section is to prove the following estimate for $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[\tau_\delta < T]$.

Theorem 3.6 *Assume that the dimension $d \geq 3$. Then, one can choose $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ in such a way that there exist constants $C, \gamma > 0$ for which*

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[\,\tau_\delta < T] \leq C\delta^\gamma T, \quad \forall\,\delta \in (0,1],\, T \geq 1,\, (\mathbf{x},\mathbf{k}) \in \mathcal{A}(M). \qquad (3.30)$$

Proof. We obviously have

$$[\,\tau_\delta < T] = [\,U_\delta \leq \tau_\delta,\, U_\delta < T] \cup [\,S_\delta \leq \tau_\delta,\, S_\delta < T] \qquad (3.31)$$

with the stopping times S_δ and U_δ defined in (3.3) and (3.4). Let us denote the first and second event appearing on the right hand side of (3.31) by $A(\delta)$ and $B(\delta)$ respectively. To show that (3.31) holds we prove that the $R^{(\delta)}_{\mathbf{x},\mathbf{k}}$ probabilities of both events can be estimated by $C\delta^\gamma T$ for some $C, \gamma > 0$: see (3.39), (3.40) and (3.44).

3.8.1 An estimate of $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A(\delta)]$

The first step towards obtaining the desired estimate will be to replace the event $A(\delta)$ whose definition involves a stopping time by an event $C(\delta)$ whose definition depends only on deterministic times, see (3.32) below. Next we use the estimate (3.21) of Proposition 3.4 for an appropriately chosen function G to reduce the question of bounding the $R^{(\delta)}_{\mathbf{x},\mathbf{k}}$ probability of $\tilde{A}(\delta)$ by an easier problem of estimating its $\mathfrak{Q}_{\mathbf{x},\mathbf{k}}$ probability ($\mathfrak{Q}_{\mathbf{x},\mathbf{k}}$ corresponds to a degenerate diffusion determined by (3.19)). The latter is achieved by using bounds on heat kernels corresponding to hypoelliptic diffusions due to Kusuoka and Stroock.

We assume in this section to simplify the notation and without any loss of generality that $h^*(M)=1$. Note that then

$$A(\delta)\subset \tilde{A}(\delta):=\left[\left|X\left(\frac{j}{q}\right)-X\left(\frac{i}{q}\right)\right|\le \frac{3}{q}:\, 1\le i\le j\le [Tq],\quad |i-j|\ge \frac{q}{p}\right] \tag{3.32}$$

and thus

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A(\delta)]\le [Tq]^2\max\left\{R^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\left|X\left(\frac{j}{q}\right)-X\left(\frac{i}{q}\right)\right|\le \frac{3}{q}\right]:\, 1\le i\le j\le [Tq],\ |i-j|\ge \frac{q}{p}\right\}. \tag{3.33}$$

Suppose that $f^{(\delta)}:\mathbb{R}^d\to[0,1]$ is a C^∞–regular function that satisfies $f(\mathbf{x})=1$, if $|\mathbf{x}|\le 4h/q$ and $f^{(\delta)}(\mathbf{x})=0$, if $|\mathbf{x}|\ge 5/q$. We assume furthermore that i,j are positive integers such that $(j-i)/q\in[0,1]$ and $\|f^{(\delta)}\|_3\le 2q^3$. For any $\mathbf{x}_0\in\mathbb{R}^d$ and $i/q\le t\le j/q$ define

$$G_j(t,\mathbf{x},\mathbf{k};\mathbf{x}_0):=\mathfrak{M}_{\mathbf{x},\mathbf{k}}f^{(\delta)}\left(X\left(\frac{j}{q}-t\right)-\mathbf{x}_0\right).$$

Obviously, we have

$$\partial_t G_j(t,\mathbf{x},\mathbf{k};\mathbf{x}_0)+\tilde{\mathcal{L}}G_j(t,\mathbf{x},\mathbf{k};\mathbf{x}_0)=0.$$

Hence, using Proposition 3.4 with $v=j/q$ and $t=i/q$ (note that $v-t\ge 1/p\ge\delta^{\epsilon_2}$ and $\epsilon_2\in(0,1/2)$), we obtain that there exists $\gamma_1>0$ such that

$$\begin{aligned}&\left|E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[f^{(\delta)}\left(X\left(\frac{j}{q}\right)-\mathbf{x}_0\right)-G_j\left(\frac{i}{q},X\left(\frac{i}{q}\right),K\left(\frac{i}{q}\right);\mathbf{x}_0\right)\,\middle|\,\mathcal{M}^{i/q}\right]\right|\\ &\le C\,\frac{j-i}{q}\,\|G_j(\cdot,\cdot,\cdot;\mathbf{x}_0)\|^{[i/q,j/q]}_{1,1,3}T^2\delta^{\gamma_1},\quad \forall\,\delta\in(0,1].\end{aligned} \tag{3.34}$$

According to [13] Theorem 2.58, p. 53 we have

$$\|G_j(\cdot,\cdot,\cdot;\mathbf{x}_0)\|^{[i/q,j/q]}_{1,1,3}\le C\|f^{(\delta)}\|_3\le Cq^3\le C\delta^{-3(\epsilon_2+\epsilon_3)},\quad j\in\{0,\dots,[qT]\}. \tag{3.35}$$

Hence combining (3.34) and (3.35) we obtain that the left hand side of (3.34) is less than, or equal to $C\,\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}$ for all $\delta\in(0,1]$. Let now $i_0=j-\dfrac{q}{p}$ so that $1\le i\le i_0\le j\le[Tq]$. We have

$$\begin{aligned}R^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\left|X\left(\frac{j}{q}\right)-X\left(\frac{i}{q}\right)\right|\le\frac{3}{q}\right]&\le E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[f^{(\delta)}\left(X\left(\frac{j}{q}\right)-X\left(\frac{i}{q}\right)\right)\right]\\ &=E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[f^{(\delta)}\left(X\left(\frac{j}{q}\right)-\mathbf{y}\right)\,\middle|\,\mathcal{M}^{i_0/q}\right]\Bigg|_{\mathbf{y}=X(i/q)}\right].\end{aligned} \tag{3.36}$$

According to (3.34) and (3.35) we can estimate the utmost right hand side of (3.36) by

$$\sup_{\mathbf{x},\mathbf{y},\mathbf{k}}\left\{\mathfrak{M}_{\mathbf{x},\mathbf{k}}f^{(\delta)}\left(X\left(\frac{1}{p}\right)-\mathbf{y}\right):\mathbf{x},\mathbf{y}\in\mathbb{R}^d,\ \mathbf{k}\in A(2M)\right\}+C\,\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}T^2. \tag{3.37}$$

To estimate the first term in (3.37) we use the following.

Lemma 3.7 *Let p, q be as in (3.2). Then, there exist positive constants C_1, C_2 and C_3 such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $\mathbf{k} \in A(2M)$, $j \in \{1, \ldots, [pT]\}$, $\delta \in (0, 1]$ we have*

$$\mathfrak{Q}_{\mathbf{x},\mathbf{k}}\left[\left|X\left(\frac{j}{p}\right) - \mathbf{y}\right| \leq \frac{5}{q}\right] \leq C_1\left(\frac{p^{C_2}}{q^d} + e^{-C_3 p}\right). \tag{3.38}$$

We postpone the proof of the lemma for a moment in order to finish the estimate of $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A(\delta)]$. Using (3.38) we obtain that the expression in (3.37) can be estimated by

$$C_1\left(\frac{p^{C_2}}{q^d} + e^{-C_3 p}\right) + C\,\delta^{\gamma_1 - 3(\epsilon_2+\epsilon_3)} T^2 \leq C_1 \delta^{(d-C_2)\epsilon_2 + d\epsilon_3} + \exp\left\{-C_3\delta^{-\epsilon_2}\right\} + C\,\delta^{\gamma_1 - 3(\epsilon_2+\epsilon_3)} T^2.$$

Hence, from (3.33), we obtain that

$$\begin{aligned} R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A(\delta)] &\leq [Tq]^2\left(C_1\delta^{(d-C_2)\epsilon_2+d\epsilon_3} + \exp\left\{-C_3\delta^{-\epsilon_2}\right\} + C\,\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}T^2\right) \\ &\leq CT^2\left(\delta^{(d-2-C_1)\epsilon_2+(d-2)\epsilon_3} + \delta^{-2(\epsilon_2+\epsilon_3)}\exp\left\{-C_3\delta^{-\epsilon_2}\right\} + \delta^{\gamma_1-5(\epsilon_2+\epsilon_3)}T^2\right) \leq C\delta^{\gamma_2}T^4 \end{aligned} \tag{3.39}$$

for $\gamma_2 := \min[(d-2-C_1)\epsilon_2 + (d-2)\epsilon_3, \gamma_1 - 5(\epsilon_2+\epsilon_3)] > 0$, provided that $\epsilon_2 + \epsilon_3 < \gamma_1/5$ and $\epsilon_2 \in (0, (d-2)\epsilon_3/(C_1+2-d))$. Here with no loss of generality we have assumed that $C_1 + 2 > d$. Recall also that $d \geq 3$. Now suppose that $\gamma_3 \in (0, \gamma_2)$. Consider two cases: $T^3 < \delta^{-\gamma_3}$ and $T^3 \geq \delta^{-\gamma_3}$. In the first one, the utmost right hand side of (3.39) can be bound from above by $C\delta^{\gamma_2-\gamma_3}T$. In the second we have a trivial bound of the left side by $\delta^{\gamma_3/3}T$. We have proved therefore that

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[A(\delta)] \leq C\delta^{\gamma}T \tag{3.40}$$

for some $C, \gamma > 0$ independent of δ and T.

The proof of Lemma 3.7. We prove this lemma by induction on j. First, we verify it for $j = 1$. Without any loss of generality we may suppose that $\mathbf{k} = (k_1, \ldots, k_d)$ and $k_d > (4dM_\delta)^{-1}$. Let $\tilde{D}_{mn} : \mathbb{R}^{d-1} \to \mathbb{R}$, $m, n = 1, \ldots, d-1$, $\tilde{E}_m : \mathbb{R}^{d-1} \to \mathbb{R}$, $m = 1, \ldots, d$ be given by

$$\tilde{D}_{pq}(\mathbf{l}) := D_{pq}(k^{-1}\mathbf{l}, k^{-1}\sqrt{k^2 - l^2}, k), \quad \tilde{E}_p(\mathbf{l}) := E_p(k^{-1}\mathbf{l}, k^{-1}\sqrt{k^2-l^2}, k),$$

when $\mathbf{l} \in Z := [\mathbf{l} \in \mathbb{B}^{d-1}_k : k^{-1}\sqrt{k^2-l^2} > (4dM_\delta)^{-1}]$, $l = |\mathbf{l}|$. These functions are C^∞ smooth and bounded together with all their derivatives. Note also that the matrix $\tilde{\mathbf{D}} = [\tilde{D}_{mn}]$ is symmetric and $\tilde{\mathbf{D}}\xi \cdot \xi \geq \lambda_0|\xi|^2$ for all $\xi \in \mathbb{R}^{d-1}$ and a certain $\lambda_0 > 0$. The projection $K(t) = (K_1(t), \ldots, K_d(t))$ of the canonical path process $(X(t;\pi), K(t;\pi))$ considered over the probability space $(\mathcal{C}, \mathcal{M}, \mathfrak{Q}_{\mathbf{x},\mathbf{k}_0})$, where $\mathbf{k}_0 := (\mathbf{l}, \sqrt{k^2-l^2})$, with $\mathbf{l} \in Z$, is a diffusion whose generator equals $\mathcal{L}$, see (3.18). It can be easily seen that $(K_1(t), \ldots, K_{d-1}(t))_{t\geq 0}$, is then a diffusion starting at $\mathbf{l}$, whose generator $\mathcal{N}$ is of the form

$$\mathcal{N}F(\mathbf{l}, \mathbf{x}) := \sum_{p=1}^{d-1} X_p^2 F(\mathbf{l}) + \sum_{q=1}^{d-1} a_q(\mathbf{l})\partial_{l_q}F(\mathbf{l}), \quad F \in C_0^\infty(\mathbb{R}^{d-1}), \tag{3.41}$$

where $a_q(\mathbf{l})$, $q = 1, \ldots, d-1$ are certain C^∞-functions and

$$X_p(\mathbf{l}) := \sum_{q=1}^{d-1} \tilde{D}^{1/2}_{pq}(\mathbf{l})\partial_{l_q}, \quad p = 1, \ldots, d-1.$$

The $(d-1)\times(d-1)$ matrix $[\tilde{D}^{1/2}_{pq}(\mathbf{l})]$ is non-degenerate when $\mathbf{l} \in Z$. Let

$$\tilde{\mathcal{N}}F(\mathbf{l}, \mathbf{x}) := \sum_{p=1}^{d-1} \tilde{X}_p^2 F(\mathbf{l}, \mathbf{x}) + \tilde{X}_0 F(\mathbf{l}, \mathbf{x}), \quad F \in C_0^\infty(\mathbb{R}^{d-1} \times \mathbb{R}^d),$$

where $\tilde{X}_0$ is a C^∞–smooth extension of the field

$$X_0(\mathbf{l}) := \frac{H_0'(k)}{k}\sum_{q=1}^{d-1} l_q\partial_{x_q} + \frac{H_0'(k)}{k}\sqrt{k^2-l^2}\,\partial_{x_d} + \sum_{q=1}^{d-1} a_q(\mathbf{l})\partial_{l_q}, \quad \mathbf{l}\in Z.$$

It can be shown, by the same type of argument as that given on pp. 122-123 of [1], that for each $(\mathbf{x},\mathbf{l})$, with $\mathbf{l}\in Z$, the linear space spanned at that point by the fields belonging to the Lie algebra generated by $X_0,\ldots,X_{d-1}$ is of dimension $2d-1$. One can also guarantee that the extensions $\tilde{X}_0,\ldots,\tilde{X}_{d-1}$ satisfy the same condition. We shall denote the respective extension of $\mathcal{N}$ by the same symbol.

Set $\mathbf{l}_0 := (k_1,\ldots,k_{d-1})$. Let $\mathcal{R}_{\mathbf{l}_0}$, $\tilde{\mathcal{R}}_{\mathbf{x},\mathbf{l}_0}$ be the path measures supported on $\mathcal{C}^{d-1}$ and $\mathcal{C}^{d,d-1}$ respectively that solve the martingale problems corresponding to the generators $\mathcal{N}$ and $\tilde{\mathcal{N}}$ with the respective initial conditions at $t=0$ given by $\mathbf{l}_0$ and $(\mathbf{x},\mathbf{l}_0)$. Let $r(t,\mathbf{x}-\mathbf{y},\mathbf{l}_1,\mathbf{l}_2)$, $t\in(0,+\infty)$, $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$, $\mathbf{l}_1,\mathbf{l}_2\in\mathbb{R}^{d-1}$ be the transition of probability density that corresponds to $\tilde{\mathcal{R}}_{\mathbf{x},\mathbf{l}_0}$. Using Corollary 3.25 p. 22 of [10] we have that for some constants $C,m>0$

$$r\,(t,\mathbf{y},\mathbf{k},\mathbf{l}) \le Ct^{-m}, \quad \forall\,\mathbf{y}\in\mathbb{R}^d,\ \mathbf{k},\ \mathbf{l}\in\mathbb{R}^{d-1},\ t\in(0,1]. \tag{3.42}$$

Denote by $\tau_Z(\pi)$ the exit time of a path $\pi\in\mathcal{C}^{d-1}$ from the set Z. For any $\pi\in\mathcal{C}^{d,d-1}$ we set also $\tilde{\tau}_Z(\pi)=\tau_Z(K(\cdot;\pi))$. Let $S:\mathbb{B}_k^{d-1}\to\mathbb{S}_k^{d-1}$ be given by

$$S(\mathbf{l}) := (l_1,\ldots,l_{d-1},\sqrt{k^2-l^2}), \quad \mathbf{l}=(l_1,\ldots,l_{d-1})\in\mathbb{B}_k^{d-1},\ l:=|\mathbf{l}|$$

and let $\tilde{S}:\mathcal{C}^{d,d-1}\to\mathcal{C}$ be given by $\tilde{S}(\pi)(t) := (X(t;\pi), S\circ K(t;\pi))$, $t\ge 0$. For any $A\in\mathcal{M}^{\tilde{\tau}_Z}$ we have $\tilde{\mathcal{R}}_{\mathbf{x},\mathbf{l}_0}[\tilde{S}^{-1}(A)] = \mathfrak{Q}_{\mathbf{x},S(\mathbf{l}_0)}[A]$. Since the event $[|X\,(1/p)-\mathbf{y}|\le 5/q]\cap[\tilde{\tau}_Z\ge 1/p]$ is $\mathcal{M}^{\tilde{\tau}_Z}$–measurable we have

$$\mathfrak{Q}_{\mathbf{x},\mathbf{k}}\left[\left|X\left(\frac{1}{p}\right)-\mathbf{y}\right|\le\frac{5}{q}\right] \le \tilde{\mathcal{R}}_{\mathbf{x},\mathbf{l}_0}\left[\left|X\left(\frac{1}{p}\right)-\mathbf{y}\right|\le\frac{5}{q},\tilde{\tau}_Z\ge\frac{1}{p}\right] + \mathcal{R}_{\mathbf{l}_0}\left[\tau_Z<\frac{1}{p}\right] \tag{3.43}$$
$$\le C\bar{\omega}_d p^m\left(\frac{4}{q}\right)^d + Ce^{-C_3p}.$$

Here $\bar{\omega}_d$ denotes the volume of $\mathbb{B}^d$. To obtain the last inequality we have used (3.42) and an estimate for non-degenerate diffusions stating that $\mathcal{R}_{\mathbf{l}_0}[\tau_Z<1/p] < Ce^{-C_3p}$ for some constants $C,C_3>0$ depending only on d and λ_0, see e.g. (2.1) p. 87 of [14]. Inequality (3.43) implies easily (3.38) for $j=1$ with $C_1=m$. To finish the induction argument assume that (3.38) holds for a certain j. We show that it holds for $j+1$ with the same constants C_1,C_2 and $C_3>0$. The latter follows easily from the Chapman-Kolmogorov equation, since

$$\mathfrak{Q}_{\mathbf{x},\mathbf{k}}\left[\left|X\left(\frac{j+1}{p}\right)-\mathbf{y}\right|\le\frac{5}{q}\right] = \iint_{\mathbb{R}^d\times\mathbb{S}_k^{d-1}} \mathfrak{Q}_{\mathbf{y},\mathbf{l}}\left[\left|X\left(\frac{j}{p}\right)-\mathbf{y}\right|\le\frac{5}{q}\right] Q\left(\frac{1}{p},\mathbf{x},\mathbf{k},d\mathbf{y},d\mathbf{l}\right)$$
$$\overset{\text{induction assumpt.}}{\le} C_1\left[\frac{p^{C_2}}{q^d}+e^{-C_3p}\right]\iint Q\left(\frac{1}{p},\mathbf{x},\mathbf{k},d\mathbf{y},d\mathbf{l}\right) = C_1\left[\frac{p^{C_2}}{q^d}+e^{-C_3p}\right]$$

and the formula (3.38) for $j+1$ follows. Here $Q(t,\mathbf{x},\mathbf{k},\cdot,\cdot)$ is the transition of probability corresponding to the path measure $\mathfrak{Q}_{\mathbf{x},\mathbf{k}}$. $\square$

3.8.2 An estimate of $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[B(\delta)]$

We start with a simple observation concerning the Hölder regularity of the K component of any path $\pi \in B(\delta)$. Let us denote $\rho := 2M_\delta^{-1}N^{-1/2}$ and

$$D := \left[\pi \in \mathcal{C}(T,\delta) : |K(t) - K(s)| \geq \rho \text{ for some } k \text{ s.t. } t_k^{(p)} \leq T \text{ and } t \in [t_k^{(p)}, t_{k+1}^{(p)}],\ s \in [t_{k-1}^{(p)}, t_k^{(p)}]\right],$$

where M_δ has been defined in (2.7) and N in (3.2). Suppose that $\pi \in B(\delta)$, then we can find $t \in [t_k^{(p)}, t_{k+1}^{(p)}]$, $s \in [t_{k-1}^{(p)}, t_k^{(p)}]$ for which $\hat{K}(t) \cdot \hat{K}(s) \leq 1 - 1/N$. This, however, implies that

$$|K(t) - K(s)|^2 \geq \frac{1}{M_\delta^2}\,|\hat{K}(t) - \hat{K}(s)|^2 \geq \frac{2}{M_\delta^2 N},$$

thus $\pi \in D$. Hence the desired estimate of $R^{(\delta)}_{\mathbf{x},\mathbf{k}}[B(\delta)]$ follows from the following lemma.

Lemma 3.8 *Under the assumptions of Theorem 3.6 there exist $C, \gamma > 0$ such that*

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[D] \leq CT\delta^\gamma, \quad \forall\, \delta \in (0,1],\ T \geq 1,\ (\mathbf{x},\mathbf{k}) \in \mathcal{A}(M). \tag{3.44}$$

Proof. We define the following events:

$$F_1 := \left[|K(t) - K(s)| \geq \rho \quad \text{for some } s,t \in [0,T],\ 0 < t - s < \frac{2}{p},\ t \leq \tau_\delta\right],$$
$$F_2 := \left[|K(t) - K(s)| \geq \rho \quad \text{for some } s,t \in [0,T],\ 0 < t - s < \frac{2}{p},\ s \geq \tau_\delta\right],$$
$$F_3 := \left[|K(\tau_\delta) - K(s)| \geq \frac{\rho}{2} \quad \text{for some } s \in [0,T],\ 0 < \tau_\delta - s < \frac{2}{p},\ \tau_\delta \leq T\right],$$
$$F_4 := \left[|K(\tau_\delta) - K(t)| \geq \frac{\rho}{2} \quad \text{for some } t \in [0,T],\ 0 < t - \tau_\delta < \frac{2}{p}\right].$$

Observe that $D \subset \bigcup_{i=1}^4 F_i$. Note that F_1, F_3 are $\mathcal{M}^{\tau_\delta}$–measurable, hence

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[F_i] = \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[F_i], \quad i = 1,3. \tag{3.45}$$

On the other hand for $i = 2,4$ we have

$$R^{(\delta)}_{\mathbf{x},\mathbf{k}}[F_i] = \int \mathfrak{Q}_{X(\tau_\delta(\pi)),K(\tau_\delta(\pi))}[F_{i,\pi}]\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}(d\pi),$$

where for a given $\pi \in \mathcal{C}$

$$F_{2,\pi} := \left[|K(t) - K(s)| \geq \rho \quad \text{for some } s,t \in [0,(T - \tau_\delta(\pi)) \wedge 0],\ 0 < t - s < \frac{2}{p}\right],$$
$$F_{4,\pi} := \left[|K(0) - K(t)| \geq \frac{\rho}{2} \quad \text{for some } t \in [0,(T - \tau_\delta(\pi)) \wedge 0],\ 0 < t < \frac{2}{p}\right].$$

Since all F_i, $i = 1,3$ and $F_{i,\pi}$, $i = 2,4$, $\pi \in \mathcal{C}$ are contained in the event

$$F := \left[|K(t) - K(s)| \geq \frac{\rho}{2} \quad \text{for some } s,t \in [0,T],\ 0 < t - s < \frac{2}{p}\right],$$

(3.44) would follow if we show that there exist $C > 0$ and $\gamma > 0$ for which

$$\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[F] \le CT\delta^{\gamma} \text{ for all } (\mathbf{x},\mathbf{k}) \in \mathcal{A}(M) \tag{3.46}$$

and

$$\mathfrak{Q}_{\mathbf{x},\mathbf{k}}[F] \le CT\delta^{\gamma} \text{ for all } (\mathbf{x},\mathbf{k}) \in \mathcal{A}(M_\delta). \tag{3.47}$$

The estimate (3.47) follows from elementary properties of diffusions, see e.g. (2.46) p. 47 of [13]. We carry on with the proof of (3.46). The argument is analogous to the proof of Theorem 1.4.6 of [14]. Let L be a multiple of p such that $L := [\delta^{-\gamma_0'}]$, where $\gamma_0' \in (1/2, 1)$ is to be specified even further later on. Let also $s_k^{(L)} := k/L$, $k = 0, 1, \ldots$. We now define the stopping times $\tau_k(\pi)$ that determine the times at which the K component of the path π performs k–th oscillation of size $\rho/8$. Let $\tau_0(\pi) := 0$ and for any $k \ge 0$

$$\tau_{k+1}(\pi) := \inf\left[s_k^{(L)} \ge \tau_k(\pi) : |K(s_k^{(L)}) - K(\tau_k(\pi))| \ge \frac{\rho}{8}\right],$$

with the convention that $\tau_{n+1} = +\infty$ when $\tau_n = +\infty$, or when the respective event is impossible. Let $N_\# := \min[n : \tau_{n+1} > T]$ and $\delta^* := \min[\tau_n - \tau_{n-1} : n = 1, \ldots, N_\#]$. Then, for a sufficiently small δ_0 and $\delta \in (0, \delta_0)$ we have $F \subset [\delta^* \le 1/p]$ so we only need to estimate $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}$ probability of the latter event.

Let $f : \mathbb{R}^d \to [0,1]$ be a function of $C_0^\infty(\mathbb{R}^d)$ class such that $f(\mathbf{0}) \equiv 1$, when $|\mathbf{k}| \le \rho/16$ and $f(\mathbf{k}) \equiv 0$, when $|\mathbf{k}| \ge \rho/8$. Let also $f_{\mathbf{l}}(\cdot) := f(\cdot - \mathbf{l})$ for any $\mathbf{l} \in \mathbb{R}^d$. Note that according to Lemma 3.5 we can choose constants A_ρ, $C > 0$, where C is independent of ρ, in such a way that $A_\rho < CT^2\rho^{-3}$ and the random sequence

$$S_N^{\mathbf{l}} := \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[f_{\mathbf{l}}\left(K\left(\frac{N+1}{L}\right)\right) \,\middle|\, \mathcal{M}^{N/L}\right] + A_\rho \frac{N}{L}, \quad N \ge 0 \tag{3.48}$$

is a $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}$–submartingale with respect to the filtration $\left(\mathcal{M}^{N/L}\right)_{N\ge0}$ for all $\mathbf{l}$ with $|\mathbf{l}| \in ((3M_\delta)^{-1}, 3M_\delta)$ provided that δ is sufficiently small. We can decompose

$$\begin{aligned}\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\delta^* \le \frac{2}{p}\right] &\le \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\delta^* \le \frac{2}{p},\, N_\# \le [\delta^{-\alpha}]\right] + \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\delta^* \le \frac{2}{p},\, N_\# > [\delta^{-\alpha}]\right] \\ &\le \sum_{i=1}^{[\delta^{-\alpha}]} \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\tau_i - \tau_{i-1} \le \frac{2}{p}\right] + \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[N_\# > [\delta^{-\alpha}]],\end{aligned} \tag{3.49}$$

where $\alpha > 0$ is to be determined later. We will show that

$$\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[N_\# > [\delta^{-\alpha}]] \le Ce^T\left(1 - \frac{\delta^{1/2(\epsilon_1+\epsilon_2)}}{2}\right)^{\delta^{-\alpha}} \tag{3.50}$$

and

$$\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\tau_{n+1} - \tau_n \le [L\delta_2^{\epsilon}]/L \,\middle|\, \mathcal{M}^{\tau_n}\right] \le C\delta^{\gamma}T^2, \tag{3.51}$$

for $0 < \gamma < \min[\epsilon_2 - 3\epsilon_1/2, \gamma_0' - (1+\epsilon_1)/2]$. From (3.48), (3.49) (3.50) and (3.51) we further conclude that

$$\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\delta^* \le \frac{1}{p}\right] \le CT^2\delta^{\gamma-\alpha} + Ce^T\left(1 - \frac{\delta^{1/2(\epsilon_1+\epsilon_2)}}{2}\right)^{\delta^{-\alpha}} \tag{3.52}$$

for some $C > 0$, independent of $\delta \in (0, 1]$ and $T \geq T_0$, provided that we choose $\alpha \in (1/2(\epsilon_1 + \epsilon_2), \gamma)$. This is possible if $\min[\epsilon_2 - 3\epsilon_1/2, \gamma_0' - (1+\epsilon_1)/2] > (\epsilon_1 + \epsilon_2)/2$, which is true if we assume $\epsilon_2 > 10\epsilon_1 > 0$ and $1 > \gamma_0' > (1+\epsilon_2)/2 + \epsilon_1$. Now, by the argument made after (3.39) we can always replace the first term on the right side of (3.52) by $CT\delta^{\gamma_1}$. We can also assume that the second term on the right hand side of (3.52) is less than or equal to $CT\delta^{\gamma_1}$. This can be seen as follows. Let $\beta := \alpha - 1/2(\epsilon_1 + \epsilon_2)$. The term in question is bounded by $C \exp\left\{T - C_1\delta^{-\beta}\right\}$ with $C_1 := \inf_{\rho \in (0,1]} \rho^{-1} \log\left(1 - \rho/2\right)^{-1}$. For $\delta^{-\beta} \geq 2T/C_1$ we get that $\exp\left\{T - C_1\delta^{-\beta}\right\}$ is less than or equal to $\exp\left\{-C_1\delta^{-\beta}/2\right\}$, while for $\delta^{-\beta} < 2T/C_1$ the left side of (3.52) is obviously less than $2T\delta^{\beta}/C_1$. In both cases we can find a bound as claimed. This proves (3.46) and hence the proof of Lemma 3.8 will be complete if we prove (3.50) and (3.51).

To this end, let $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}$, $\pi \in \mathcal{C}$ denote the family of the regular conditional probability distributions that corresponds to $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[\,\cdot \mid \mathcal{M}^{\tau_n}]$. Then, there exists a $\mathcal{M}^{\tau_n}$ measurable, null $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}$ probability event Z such that for each $\pi \notin Z$ and each $\mathbf{l} \in \mathbb{R}^d_*$ the random sequence

$$S^{\mathbf{l}}_{N,\pi} := S^{\mathbf{l}}_N \mathbf{1}_{[0,N/L]}(\tau_n(\pi)), \quad N \geq 0$$

is an $\left(\mathcal{M}^{N/L}\right)_{N\geq 0}$ submartingale under $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}$. Let $T_{n,\pi} := \tau_{n+1} \wedge (\tau_n(\pi) + 2[L\delta^{\epsilon}]/L)$, where $\epsilon \in (0, 1)$ is a constant to be chosen later on. We can choose the event Z in such a way that

$$\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}[T_{n,\pi} \geq \tau_n(\pi)] = 1, \quad \forall\, \pi \notin Z. \tag{3.53}$$

Let $\tilde{S}_{N,\pi} := S^{K(\tau_n(\pi))}_{N,\pi}$, then the submartingale property of $\left(\tilde{S}_{N,\pi}\right)_{N\geq 0}$ and (3.53) imply that

$$\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi} \tilde{S}_{LT_{n,\pi},\pi} \geq \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi} \tilde{S}_{L\tau_n(\pi),\pi} = 1 + A_\rho \tau_n(\pi), \tag{3.54}$$

provided that $\gamma_0 \geq (1 + \epsilon_1)/2$. The latter condition assures that $\rho \geq C/(L\sqrt{\delta})$ so that K does not change by more than ρ during the time $1/L$. In consequence of (3.54) we have

$$\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}\left[f_{K(\tau_n(\pi))}\left(K\left(T_{n,\pi} + \frac{1}{L}\right)\right)\right] + 2A_\rho\delta^{\epsilon} \geq 1, \tag{3.55}$$

as $T_{n,\pi} - \tau_n(\pi) \leq 2[L\delta^{\varepsilon}]/L$. Since

$$\left| f_{K(\tau_n(\pi))}\left(K\left(T_{n,\pi} + \frac{1}{L}\right)\right) - f_{K(\tau_n(\pi))}\left(K\left(T_{n,\pi}\right)\right)\right| \leq \frac{C}{L\rho\delta^{1/2}}$$

we obtain from (3.55)

$$2A_\rho\delta^{\epsilon} \geq \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}\left[1 - f_{K(\tau_n(\pi))}\left(K\left(T_{n,\pi}\right)\right)\right] - \frac{C}{L\rho\delta^{1/2}}$$

so in particular

$$\begin{aligned} 2A_\rho\delta^{\epsilon} + \frac{C}{L\rho\delta^{1/2}} &\geq \tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}\left[1 - f_{K(\tau_n(\pi))}\left(K\left(\tau_{n+1}\right)\right),\, \tau_{n+1} \leq \tau_n(\pi) + \frac{[L\delta^{\epsilon}]}{L}\right] \\ &= \tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k},\pi}\left[\tau_{n+1} \leq \tau_n(\pi) + \frac{[L\delta^{\epsilon}]}{L}\right]. \end{aligned} \tag{3.56}$$

We have shown, therefore, that

$$\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\tau_{n+1} - \tau_n \leq \frac{[L\delta^{\epsilon}]}{L} \,\middle|\, \mathcal{M}^{\tau_n}\right] \leq \frac{CT^2\delta^{\epsilon}}{\rho^3} + \frac{C}{L\rho\delta^{1/2}} \leq C(\delta^{\epsilon - 3\epsilon_1/2}T^2 + \delta^{\gamma_0' - (1+\epsilon_1)/2}) \leq C\delta^{\gamma_1}T^2 \tag{3.57}$$

for $\gamma_1 < \min[\epsilon - 3\epsilon_1/2, \gamma_0' - (1+\epsilon_1)/2]$ and some constant $C > 0$. We can always assume that $T^2\delta^{\gamma_1/2} \le 1$. If otherwise, we can always write $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[F] \le T\delta^{\gamma/4}$ and (3.46) follows. In particular, selecting $\epsilon := (\epsilon_1+\epsilon_2)/2$, one concludes from (3.57) that

$$\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}[\exp\{-(\tau_{n+1}-\tau_n)\}|\mathcal{M}^{\tau_n}] \le e^{-\delta^{(\epsilon_1+\epsilon_2)/2}}\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\tau_{n+1}-\tau_n \ge \frac{[L\delta^{(\epsilon_1+\epsilon_2)/2}]}{L}\,\middle|\,\mathcal{M}^{\tau_n}\right]$$

$$+\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\tau_{n+1}-\tau_n \le \frac{[L\delta^{(\epsilon_1+\epsilon_2)/2}]}{L}\,\middle|\,\mathcal{M}^{\tau_n}\right] \stackrel{(3.57)}{\le} e^{-\delta^{(\epsilon_1+\epsilon_2)/2}} + C\left(1-e^{-\delta^{(\epsilon_1+\epsilon_2)/2}}\right)\delta^{\gamma/2}$$

$$< 1 - \frac{\delta^{(\epsilon_1+\epsilon_2)/2}}{2} \tag{3.58}$$

provided that δ is sufficiently small. From (3.58) one concludes easily, see e.g. Lemma 1.4.5 p. 38 of [14], that (3.50) holds.

On the other hand, taking $\epsilon = \epsilon_2$ in (3.57) we obtain (3.51) with $0 < \gamma < \min[\epsilon_2 - 3\epsilon_1/2, \gamma_0' - (1+\epsilon_1)/2]$. Hence the proof of Lemma 3.8 is now complete. □

3.9 The estimation of the convergence rate. The proof of Theorem 2.1.

Recall that ϕ_δ, $\bar{\phi}$ satisfy (2.8), (2.10), respectively, with the initial condition ϕ_0. We start with the following lemma.

Lemma 3.9 *Assume that ϕ_0 satisfies the hypotheses formulated in Section 2.5. Then,*

$$\|\bar{\phi}\|^{[0,T]}_{0,0,0} \le \|\phi_0\|_{0,0}, \quad \sum_{i=1}^d \|\partial_{x_i}\bar{\phi}\|^{[0,T]}_{0,0,0} \le \|\phi_0\|_{1,0}. \tag{3.59}$$

Furthermore, there exists a constant $C > 0$ such that for all $T \ge 1$

$$\|\partial_t\bar{\phi}\|^{[0,T]}_{0,0,0} \le C\|\phi_0\|_{1,2}. \tag{3.60}$$

In addition, for any nonnegative integer valued multi-index $\gamma = (\alpha_1,\alpha_2,\alpha_3)$ satisfying $|\gamma| \le 3$ we have

$$\sum_{i_1,i_2,i_3=1}^d \|\partial^\gamma_{k_{i_1},k_{i_2},k_{i_3}}\bar{\phi}\|^{[0,T]}_{0,0,0} \le CT^{|\gamma|}\|\phi_0\|_{1,4}, \tag{3.61}$$

Proof. The estimates (3.59) follow directly from differentiating (2.10) with respect to $\mathbf{x}$. To obtain the estimates (3.60) and (3.61) we note first that the application of the operator $\tilde{\mathcal{L}}$ to both sides of (2.10) and the maximum principle leads to the estimate $\|\tilde{\mathcal{L}}\bar{\phi}(t,\mathbf{x},\cdot)\|_{L^\infty(A(M))} \le \|\tilde{\mathcal{L}}\phi_0\|_{L^\infty(\mathcal{A}(M))}$ for all $t \ge 0$, hence we conclude bound (3.60).

In fact, thanks to already proven estimate (3.59) we conclude that $\|\mathcal{L}\bar{\phi}(t,\mathbf{x},\cdot)\|_{L^\infty(A(M))} \le C\|\phi_0\|_{1,2}$ for some constant $C > 0$ and all $(t,\mathbf{x}) \in [0,+\infty)\times\mathbb{R}^d$. Let Z be as in the proof of Lemma 3.7. Define $S : Z \times [M^{-1},M] \to A(M)$ as $S(\mathbf{l},k) := (\mathbf{l},\sqrt{k^2-l^2})$, where $l = |\mathbf{l}|$. Let also $\psi(\mathbf{l},k) = \bar{\phi}\circ S(\mathbf{l},k)$. We have $(\mathcal{L}_{\mathbf{k}}\bar{\phi})\circ S(\mathbf{l},k) = \mathcal{N}\psi(\mathbf{l},k)$, see (3.41). The L^p estimates for elliptic partial differential equations, see e.g. Theorem 9.13 p. 239 of [7] allow us to estimate

$$\|\psi\|_{W^{2,p}(Z)} \le C(\|\psi\|_{L^p(Z)} + \|\mathcal{N}\psi\|_{L^p(Z)}) \le C\|\phi_0\|_{1,2}.$$

Choosing p sufficiently large we obtain that $\sum_i \|\partial_{l_i}\psi\|_{L^\infty(Z)} \le C\|\phi_0\|_{1,2}$, which in fact implies that $\|\mathbf{D}(\cdot)\nabla_{\mathbf{k}}\bar{\phi}(t,\cdot)\|_{L^\infty(S(Z))} \le C\|\phi_0\|_{1,2}$. Obviously, one can find a covering of $A(M)$ with charts

corresponding to different choices of the components of $\mathbf{k}$ being projected onto the hyperplane $\mathbb{R}^{d-1}$ and we obtain in that way that $\|\mathbf{D}(\cdot)\nabla_{\mathbf{k}}\bar{\phi}(t,\cdot)\|_{L^\infty(\mathcal{A}(M))} \le C\|\phi_0\|_{1,2}$ for all $t \ge 0$. Since the rank of the matrix $\mathbf{D}(\hat{\mathbf{k}},k)$ equals $d-1$, with the kernel spanned by the vector $\mathbf{k}$, we obtain in that way the L^∞ estimates of directional derivatives in any direction perpendicular to $\mathbf{k}$. We still need to obtain the L^∞ bound on the derivative in the direction $\mathbf{k}$, denoted by $\partial_n := k_1\partial_{k_1} + \ldots + k_d\partial_{k_d}$. To that purpose we apply ∂_n to both sides of (2.10) and after a straightforward calculation we get $\partial_t\partial_n\bar{\phi} = \tilde{\mathcal{L}}\partial_n\bar{\phi} - 2\mathcal{L}_{\mathbf{k}}\bar{\phi} + \mathcal{L}_1\bar{\phi} + H_0''(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}\bar{\phi}$, where

$$\mathcal{L}_1\bar{\phi} := \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(\partial_k D_{mn}(\hat{\mathbf{k}},k)\frac{\partial\bar{\phi}}{\partial k_n}\right).$$

Note that $\mathbf{D}(\hat{\mathbf{k}},k)\hat{\mathbf{k}} = \mathbf{0}$ implies that $\partial_k\mathbf{D}(\hat{\mathbf{k}},k)\hat{\mathbf{k}} = \mathbf{0}$ hence $\|\mathcal{L}_1\bar{\phi}(t,\cdot)\|_{L^\infty(\mathcal{A}(M))} \le C\|\phi_0\|_{1,2}$. We already know that $\mathcal{L}_{\mathbf{k}}\bar{\phi}$ and $\|\nabla_{\mathbf{x}}\bar{\phi}\|_{L^\infty(\mathcal{A}(M))}$ are bounded, hence $\|\partial_n\bar{\phi}(t,\cdot)\|_{L^\infty(\mathcal{A}(M))} \le C\|\phi_0\|_{1,2}T$ for $t \in [0,T]$. We have shown therefore that $\|\bar{\phi}(t,\cdot)\|_{1,1} \le C\|\phi_0\|_{1,2}T$ for $t\in[0,T]$. The above procedure can be iterated in order to obtain the estimates of the suprema of derivatives of the higher order. □

Proof of Theorem 2.1. Let $u \in [\delta^{\gamma_0'},T]$, where we assume that γ_0' (as in the statement of Lemma 3.5) belongs to the interval $(1/2,1)$. Substituting for $G(t,\mathbf{x},\mathbf{k}) := \bar{\phi}(u-t,\mathbf{x},\mathbf{k})$, $\zeta\equiv 1$ into (3.21) we obtain (taking $v=u$, $t=\delta^{\gamma_0'}$)

$$\left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\phi_0(X(u),K(u)) - \bar{\phi}(u-\delta^{\gamma_0'},X(\delta^{\gamma_0'}),K(\delta^{\gamma_0'})) - \int\limits_{\delta^{\gamma_0'}}^{u}(\partial_\varrho + \widehat{\mathcal{L}}_\varrho)G(\varrho,X(\varrho),K(\varrho))\,d\varrho\right]\right|$$
$$\le C\|G\|^{[0,T]}_{1,1,3}\delta^{\gamma_1}T^2, \quad \forall\,\delta\in(0,1]. \tag{3.62}$$

Using the fact that $|X(\delta^{\gamma_0'})-\mathbf{x}| \le C\delta^{\gamma_0'}$, $|K(\delta^{\gamma_0'})-\mathbf{k}| \le C\delta^{\gamma_0'-1/2}$, $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}$–a.s. for some deterministic constant $C>0$, cf. (3.11), and Lemma 3.9 we obtain that there exist constants $C,\gamma>0$ such that

$$\left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\phi_0(X(u),K(u)) - \bar{\phi}(u,\mathbf{x},\mathbf{k}) - \int\limits_0^u(\partial_\varrho + \widehat{\mathcal{L}}_\varrho)G(\varrho,X(\varrho),K(\varrho))\,d\varrho\right]\right|$$
$$\le C\|G\|^{[0,T]}_{1,1,3}\delta^{\gamma}T^2, \quad \delta\in(0,1],\ T\ge 1,\ u\in[0,T]. \tag{3.63}$$

We have however

$$\left|E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\phi_0(X(u),K(u)) - \bar{\phi}(u,\mathbf{x},\mathbf{k}),\ \tau_\delta\ge T\right]\right| = \left|\tilde{E}^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\phi_0(X(u),K(u)) - \bar{\phi}(u,\mathbf{x},\mathbf{k}),\ \tau_\delta\ge T\right]\right|$$
$$\stackrel{(3.63)}{\le} C\|G\|^{[0,T]}_{1,1,3}\delta^{\gamma}T^2 + \left(2\|\phi_0\|_{0,0} + T\|G\|^{[0,T]}_{1,1,2}\right)\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[\tau_\delta<T]. \tag{3.64}$$

Using $\mathcal{M}^{\tau_\delta}$ measurability of the event $[\tau_\delta<T]$ we obtain that $\tilde{Q}^{(\delta)}_{\mathbf{x},\mathbf{k}}[\tau_\delta<T] = R^{(\delta)}_{\mathbf{x},\mathbf{k}}[\tau_\delta<T]$ and by virtue of Theorem 3.6 we can estimate the right hand side of (3.64) by

$$C\|G\|^{[0,T]}_{1,1,3}\delta^{\gamma}T^2 + C\delta^{\gamma}T\left(2\|\phi_0\|_{0,0} + T\|G\|^{[0,T]}_{1,1,2}\right) \stackrel{\text{Lemma 3.9}}{\le} C\delta^{\gamma}T^5.$$

On the other hand, the expression under the absolute value on the utmost left hand side of (3.64) equals

$$E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\phi_0(X(u),K(u)) - \bar{\phi}(u,\mathbf{x},\mathbf{k})\right] - E^{(\delta)}_{\mathbf{x},\mathbf{k}}\left[\phi_0(X(u),K(u)) - \bar{\phi}(u,\mathbf{x},\mathbf{k}),\ \tau_\delta<T\right].$$

The second term can be estimated by

$$2\|\phi_0\|_{0,0} R^{(\delta)}_{\mathbf{x},\mathbf{k}}[\tau_\delta < T] \overset{(3.30)}{\leq} C\delta^\gamma \|\phi_0\|_{0,0} T,$$

by virtue of Theorem 3.6. Since

$$\mathbb{E}\phi_\delta\left(\frac{u}{\delta}, \frac{\mathbf{x}}{\delta}, \mathbf{k}\right) = \mathbb{E}\phi_0(\mathbf{z}^{(\delta)}(u;\mathbf{x},\mathbf{k}), \mathbf{m}^{(\delta)}(u;\mathbf{x},\mathbf{k})) = E^{(\delta)}_{\mathbf{x},\mathbf{k}}\phi_0(X(u), K(u))$$

we conclude from the above that the left hand side of (2.11) can be estimated by $C\delta^\gamma\|\phi_0\|_{1,4}T^5$ for some constants $C, \gamma > 0$ independent of $\delta > 0$, $T \geq 1$. The bound appearing on the right hand side of (2.11) can be now concluded by the same argument as the one used after (3.39). $\square$

4 Momentum diffusion to spatial diffusion: proof of Theorem 2.5

We show in this section that solutions of the momentum diffusion equation (2.10) in the long-time, large space limit converge to the solutions of the spatial diffusion equation (2.12). We first recall the setup of Theorem 2.5. Let $\bar\phi_\gamma(t,\mathbf{x},\mathbf{k}) = \bar\phi(t/\gamma^2, \mathbf{x}/\gamma, \mathbf{k})$, where $\bar\phi$ satisfies (2.10) and let $w(t,\mathbf{x},k)$ be the solution of the spatial diffusion equation (2.12). In order to prove Theorem 2.5 we need to show that the re-scaled solution $\phi_\gamma(t,\mathbf{x},\mathbf{k})$ converges as $\gamma \to 0$ in the space $C([0,T]; L^\infty(\mathcal{A}(M)))$ to $w(t,\mathbf{x},\mathbf{k})$, so that

$$\|w(t) - \bar\phi_\gamma(t)\|_{L^\infty(\mathcal{A}(M))} \leq C\left(\gamma T + \gamma^{1/2}\right)\|\phi_0\|_{2,0}, \quad 0 \leq t \leq T. \tag{4.1}$$

Proof of Theorem 2.5. The proof is quite standard. We present it for the sake of completeness and convenience to the reader. The function $\bar\phi_\gamma$ is the unique $C_b^{1,1,2}([0,+\infty), \mathbb{R}^{2d}_*)$-solution to

$$\gamma^2\frac{\partial\bar\phi_\gamma}{\partial t} = \sum_{m,n=1}^d \frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial\bar\phi_\gamma}{\partial k_n}\right) + \gamma H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}\bar\phi_\gamma. \tag{4.2}$$
$$\bar\phi_\gamma(0,\mathbf{x},k) = \phi_0(\mathbf{x},\mathbf{k}),$$

see Remark 2.3. We represent $\bar\phi_\gamma$ as

$$\bar\phi_\gamma = w + \gamma w_1 + \gamma^2 w_2 + R. \tag{4.3}$$

Here w is the solution of the diffusion equation (2.12), the correctors w_1 and w_2 will be constructed explicitly, and the remainder R will be shown to be small. The first corrector w_1 is the unique solution of zero mean over each sphere $\mathbb{S}^{d-1}_k$ of the equation

$$\sum_{m,n=1}^d \frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial w_1}{\partial k_n}\right) = -H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}w. \tag{4.4}$$

It has an explicit form

$$w_1(t,\mathbf{x},\mathbf{k}) = \sum_{j=1}^d \chi_j(\mathbf{k})\frac{\partial w(t,\mathbf{x},k)}{\partial x_j} \tag{4.5}$$

with the functions χ_j defined in (2.14). The second order corrector w_2 is the unique zero mean over each sphere $\mathbb{S}^{d-1}_k$ solution of the equation

$$\sum_{m,n=1}^d \frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial w_2}{\partial k_n}\right) = \frac{\partial w}{\partial t} - H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}w_1. \tag{4.6}$$

Note that the expression on the right hand side of (4.6) is of zero mean since thanks to (2.12) and equality (2.13) we have

$$\frac{\partial w}{\partial t} = \frac{1}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} w_1 d\Omega(\hat{\mathbf{k}}).$$

Equations (4.4) and (4.6) for various values of $k = |\mathbf{k}|$ are decoupled. As a consequence of this fact and the regularity properties for solutions of elliptic equations on a sphere we have that w_1, w_2 belong to $C([0,T]; L^\infty(\mathcal{A}(M)))$. More explicitly, we may represent the function w_2 as

$$w_2(t, \mathbf{x}, \mathbf{k}) = \sum_{j,l=1}^{d} \psi_{jl}(\mathbf{k}) \frac{\partial^2 w(t, \mathbf{x}, \mathbf{k})}{\partial x_j \partial x_l}.$$

The functions $\psi_{jm}(\mathbf{k})$ satisfy

$$\sum_{m,n=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{\mathbf{k}}, k) \frac{\partial \psi_{jl}}{\partial k_n} \right) = -H_0'(k)\hat{k}_j \chi_l(\mathbf{k}) + a_{jl}(k). \tag{4.7}$$

A unique mean-zero, bounded solution of (4.7) exists according to the Fredholm alternative combined the regularity properties for solutions of (4.7) on each sphere $\mathbb{S}_k^{d-1}$. With the above definitions of w, w_1, w_2, equation (4.2) for $\bar{\phi}_\gamma$ implies that the remainder R in (4.3) satisfies

$$\gamma^2 \frac{\partial R}{\partial t} + \gamma^3 \frac{\partial w_1}{\partial t} + \gamma^4 \frac{\partial w_2}{\partial t} - \gamma H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} R - \gamma^3 H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} w_2 = \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{\mathbf{k}}, k) \frac{\partial R}{\partial k_n} \right).$$

We re-write this equation in the form

$$\begin{aligned} &\frac{\partial R}{\partial t} - \frac{1}{\gamma} H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} R - \frac{1}{\gamma^2} \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{\mathbf{k}}, k) \frac{\partial R}{\partial k_n} \right) = f \\ &R(0, \mathbf{x}, \mathbf{k}) = \phi_0(\mathbf{x}, \mathbf{k}) - \bar{\phi}_0(\mathbf{x}, \mathbf{k}) - \gamma w_1(0, \mathbf{x}, \mathbf{k}) - \gamma^2 w_2(0, \mathbf{x}, \mathbf{k}), \end{aligned} \tag{4.8}$$

where $f := -\gamma \partial_t w_1 - \gamma^2 \partial_t w_2 - \gamma H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} w_2$. Here, as before, R is understood as the unique solution to (4.8) that belongs to $C_b^{1,1,2}([0, +\infty), \mathbb{R}_*^{2d})$. We may split $R = R_1 + R_2$ according to the initial data and forcing in the equation: R_1 satisfies

$$\begin{aligned} &\frac{\partial R_1}{\partial t} - \frac{1}{\gamma} H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} R_1 - \frac{1}{\gamma^2} \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{\mathbf{k}}, k) \frac{\partial R_1}{\partial k_n} \right) = f, \\ &R(0, \mathbf{x}, \mathbf{k}) = -\gamma w_1(0, \mathbf{x}, \mathbf{k}) - \gamma^2 w_2(0, \mathbf{x}, \mathbf{k}) \end{aligned} \tag{4.9}$$

and the initial time boundary layer corrector R_2 satisfies

$$\begin{aligned} &\frac{\partial R_2}{\partial t} - \frac{1}{\gamma} H_0'(k)\hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} R_2 - \frac{1}{\gamma^2} \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{\mathbf{k}}, k) \frac{\partial R_2}{\partial k_n} \right) = 0 \\ &R_2(0, \mathbf{x}, \mathbf{k}) = \phi_0(\mathbf{x}, \mathbf{k}) - \bar{\phi}_0(\mathbf{x}, \mathbf{k}). \end{aligned} \tag{4.10}$$

Using the probabilistic representation for the solution of (4.9) as well as the regularity of w_1 and w_2 we obtain that

$$\|R_1(t)\|_{L^\infty(\mathcal{A}(M))} \le C\gamma T, \quad 0 \le t \le T. \tag{4.11}$$

To obtain the bound for R_2 we consider $R_2^\gamma(t,\mathbf{x},\mathbf{k}) := R_2(\gamma^{3/2}t,\mathbf{x},\mathbf{k})$. This function satisfies

$$\begin{aligned}
&\frac{\partial R_2^\gamma}{\partial t} - \gamma^{1/2}H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}R_2^\gamma - \frac{1}{\gamma^{1/2}}\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial R_2^\gamma}{\partial k_n}\right) = 0\\
&R_2^\gamma(0,\mathbf{x},\mathbf{k}) = \phi_0(\mathbf{x},\mathbf{k}) - \bar{\phi}_0(\mathbf{x},\mathbf{k}).
\end{aligned}$$

We also define $\tilde{R}_2^\gamma$, the solution of

$$\begin{aligned}
&\frac{\partial \tilde{R}_2^\gamma}{\partial t} - \frac{1}{\gamma^{1/2}}\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial \tilde{R}_2^\gamma}{\partial k_n}\right) = 0\\
&\tilde{R}_2^\gamma(0,\mathbf{x},\mathbf{k}) = \phi_0(\mathbf{x},\mathbf{k}) - \bar{\phi}_0(\mathbf{x},\mathbf{k}).
\end{aligned} \tag{4.12}$$

The uniform ellipticity of the right hand side of (4.12) on each sphere $\mathbb{S}_k^{d-1}$ implies, see e.g. Proposition 13.3, p. 55 of [15] that the function $\tilde{R}_2^\gamma$ satisfies the decay estimate on each sphere

$$\|\tilde{R}_2^\gamma(t)\|_{L^\infty(\mathbb{S}^{d-1})} \le \frac{C\gamma^{(d-1)/4}}{t^{(d-1)/2}}\|\phi_0\|_{L^1(\mathbb{S}_k^{d-1})} \le \frac{C\gamma^{(d-1)/4}}{t^{(d-1)/2}}\|\phi_0\|_{L^\infty(\mathbb{S}_k^{d-1})} \tag{4.13}$$

for $t\in[0,T]$ and, similarly,

$$\|\nabla_{\mathbf{x}}\tilde{R}_2^\gamma(t)\|_{L^\infty(\mathbb{S}_k^{d-1})} \le \frac{C\gamma^{(d-1)/4}}{t^{(d-1)/2}}\|\phi_0\|_{1,0}.$$

Furthermore, the difference $q^\gamma = R_2^\gamma - \tilde{R}_2^\gamma$ satisfies

$$\begin{aligned}
&\frac{\partial q^\gamma}{\partial t} - \gamma^{1/2}H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}q^\gamma - \frac{1}{\gamma^{1/2}}\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{\mathbf{k}},k)\frac{\partial q^\gamma}{\partial k_n}\right) = \gamma^{1/2}H_0'(k)\hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}}\tilde{R}_2^\gamma\\
&q^\gamma(0,\mathbf{x},\mathbf{k}) = 0.
\end{aligned} \tag{4.14}$$

We conclude, using the probabilistic representation of the solution of (4.14), that

$$\|q^\gamma(t)\|_{L^\infty(\mathcal{A}(M))} \le C\gamma^{1/2}t\|\phi_0\|_{1,0}$$

and thus

$$\begin{aligned}
\|R_2(\gamma^{3/2})\|_{L^\infty(\mathcal{A}(M))} &\le \|R_2^\gamma(1)\|_{L^\infty(\mathcal{A}(M))} + \|q^\gamma(1)\|_{L^\infty(\mathcal{A}(M))}\\
&\le C\left(\gamma^{(d-1)/4}\|\phi_0\|_{0,0} + \gamma^2\|\phi_0\|_{1,0}\right).
\end{aligned}$$

The maximum principle for (4.10) implies that we have the above estimate for all $t \ge \gamma^{3/2}$:

$$\|R_2(t)\|_{L^\infty(\mathcal{A}(M))} \le C\left(\gamma^{(d-1)/4}\|\phi_0\|_{0,0} + \gamma^{1/2}\|\nabla_{\mathbf{x}}\phi_0\|_{1,0}\right), \quad t \ge \gamma^{3/2}. \tag{4.15}$$

Combining (4.3), (4.11) and (4.15) we conclude that

$$\|w(t) - \bar{\phi}_\gamma(t)\|_{L^\infty(\mathcal{A}(M))} \le C\left(\gamma T + \gamma^{(d-1)/4} + \gamma^{1/2}\right)\|\phi_0\|_{1,0}, \quad \gamma^{3/2} \le t \le T, \tag{4.16}$$

and thus (4.1) follows, as $d \ge 3$. This finishes the proof of Theorem 2.5. $\square$

5 The spatial diffusion of wave energy

In this section we consider an application of the previous results to the random geometrical optics regime of propagation of acoustic waves. We show that when the wave length is much shorter than the correlation length of the random medium, there exist temporal and spatial scales where the energy density of the wave undergoes the spatial diffusion. We start with the wave equation in dimension $d \geq 3$

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 \phi}{\partial t^2} - \Delta\phi = 0 \tag{5.1}$$

and assume that the wave speed has the form $c(\mathbf{x}) = c_0 + \sqrt{\delta} c_1(\mathbf{x})$. Here $c_0 > 0$ is the constant sound speed of the uniform background medium, while the small parameter $\delta \ll 1$ measures the strength of the mean zero random perturbation c_1. Rescaling the spatial and temporal variables $\mathbf{x} = \mathbf{x}'/\delta$ and $t = t'/\delta$ we obtain (after dropping the primes) equation (5.1) with a rapidly fluctuating wave speed

$$c_\delta(\mathbf{x}) = c_0 + \sqrt{\delta} c_1\left(\frac{\mathbf{x}}{\delta}\right). \tag{5.2}$$

It is convenient to rewrite (5.1) as the system of acoustic equations for the "pressure" $p = \phi_t/c$ and "acoustic velocity" $\mathbf{u} = -\nabla\phi$:

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} + \nabla\left(c_\delta(\mathbf{x})p\right) &= 0\\ \frac{\partial p}{\partial t} + c_\delta(\mathbf{x})\nabla\cdot\mathbf{u} &= 0.\end{aligned} \tag{5.3}$$

We will denote for brevity $\mathbf{v} = (\mathbf{u}, p) \in \mathbb{R}^{d+1}$ and write (5.3) in the more general form of a first order linear symmetric hyperbolic system. To do so we introduce symmetric matrices A_δ and D^j defined by

$$A_\delta(\mathbf{x}) = \mathrm{diag}(1, 1, 1, c_\delta(\mathbf{x})), \qquad \text{and} \qquad D^j = \mathbf{e}_j \otimes \mathbf{e}_{d+1} + \mathbf{e}_{d+1} \otimes \mathbf{e}_j,\ \ j = 1, \ldots, d. \tag{5.4}$$

Here $\mathbf{e}_m \in \mathbb{R}^{d+1}$ is the standard orthonormal basis: $(\mathbf{e}_m)_k = \delta_{mk}$.

We consider the initial data for (5.3) as a mixture of states. Let S be a measure space equipped with a non-negative finite measure μ. A typical example is that the initial data is random, S is the state space and μ is the corresponding probability measure. We assume that for each parameter $\zeta \in S$ and $\varepsilon, \delta > 0$ the initial data is given by $\mathbf{v}_\varepsilon^\delta(0, \mathbf{x}; \zeta) := (-\varepsilon\nabla\phi_0^\varepsilon(\mathbf{x}), 1/c_\delta(\mathbf{x})\dot{\phi}_0^\varepsilon(\mathbf{x}))$ and $\mathbf{v}_\varepsilon^\delta(t, \mathbf{x}; \zeta)$ solves the system of equations

$$\frac{\partial \mathbf{v}_\varepsilon^\delta}{\partial t} + \sum_{j=1}^d A_\delta(\mathbf{x}) D^j \frac{\partial}{\partial x^j}\left(A_\delta(\mathbf{x})\mathbf{v}_\varepsilon^\delta(\mathbf{x})\right) = 0. \tag{5.5}$$

The set of initial data is assumed to form an ε-oscillatory and compact at infinity family [5] as $\varepsilon \to 0$. By the above we mean that for any continuous, compactly supported function $\varphi : \mathbb{R}^d \to \mathbb{R}$ we have

$$\lim_{R\to+\infty} \limsup_{\varepsilon\to 0^+} \int_{|\mathbf{k}|\geq R/\varepsilon} |\widehat{\varphi \mathbf{v}_\varepsilon^\delta}|^2 d\mathbf{k} \to 0 \text{ and } \lim_{R\to+\infty} \limsup_{\varepsilon\to 0^+} \int_{|\mathbf{x}|\geq R} |\mathbf{v}_\varepsilon^\delta|^2 d\mathbf{x} \to 0$$

for a fixed realization $\zeta \in S$ of the initial data and each $\delta > 0$. In the regime of geometric acoustics the scale ε of oscillations of the initial data is much smaller than the correlation length δ of the medium: $\varepsilon \ll \delta \ll 1$.

The dispersion matrix for (5.5) is

$$P_0^\delta(\mathbf{x},\mathbf{k}) = i\sum_{j=1}^d A_\delta(\mathbf{x})k_j D^j A_\delta(\mathbf{x}) = i\sum_{j=1}^d c_\delta(\mathbf{x})k_j D^j = ic_\delta(\mathbf{x})\left(\tilde{\mathbf{k}}\otimes\mathbf{e}_{d+1}+\mathbf{e}_{d+1}\otimes\tilde{\mathbf{k}}\right), \tag{5.6}$$

where $\tilde{\mathbf{k}} = \sum_{j=1}^d k_j\mathbf{e}_j$. The self-adjoint matrix $(-iP_0^\delta)$ has an eigenvalue $H_0 = 0$ of the multiplicity $d-1$, and two simple eigenvalues

$$H_\pm^\delta(\mathbf{x},\mathbf{k}) = \pm c_\delta(\mathbf{x})|\mathbf{k}|. \tag{5.7}$$

Its eigenvectors are

$$\mathbf{b}_m^0 = \left(\mathbf{k}_m^\perp, 0\right),\ m = 1,\ldots,d-1;\ \ \mathbf{b}_\pm = \frac{1}{\sqrt{2}}\left(\frac{\tilde{\mathbf{k}}}{|\mathbf{k}|}\pm\mathbf{e}_{d+1}\right), \tag{5.8}$$

where $\mathbf{k}_m^\perp \in \mathbb{R}^d$ is the orthonormal basis of vectors orthogonal to $\mathbf{k}$.

The $(d+1)\times(d+1)$ Wigner matrix of a mixture of solutions of (5.5) is defined by

$$W_\varepsilon^\delta(t,\mathbf{x},\mathbf{k}) = \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d}\int\limits_S e^{i\mathbf{k}\cdot\mathbf{y}}\mathbf{v}_\varepsilon^\delta(t,\mathbf{x}-\frac{\varepsilon\mathbf{y}}{2};\zeta)\mathbf{v}_\varepsilon^{\delta*}(t,\mathbf{x}+\frac{\varepsilon\mathbf{y}}{2};\zeta)d\mathbf{y}\mu(d\zeta). \tag{5.9}$$

It is well-known, see [5, 11, 12], that for each fixed $\delta > 0$ (and even without introduction of a mixture of states) when $W_\varepsilon^\delta(t=0)$ converges weakly in $\mathcal{S}'(\mathbb{R}^d\times\mathbb{R}^d)$, as $\varepsilon\to 0$, to

$$W_0(\mathbf{x},\mathbf{k}) = u_+^0(\mathbf{x},\mathbf{k})\mathbf{b}_+(\mathbf{k})\otimes\mathbf{b}_+(\mathbf{k}) + u_-^0(\mathbf{x},\mathbf{k})\mathbf{b}_-(\mathbf{k})\otimes\mathbf{b}_-(\mathbf{k}). \tag{5.10}$$

then $W_\varepsilon^\delta(t)$ converges weakly in $\mathcal{S}'(\mathbb{R}^d\times\mathbb{R}^d)$ to

$$U^\delta(t,\mathbf{x},\mathbf{k}) = u_+^\delta(t,\mathbf{x},\mathbf{k})\mathbf{b}_+(\mathbf{k})\otimes\mathbf{b}_+(\mathbf{k}) + u_-^\delta(t,\mathbf{x},\mathbf{k})\mathbf{b}_-(\mathbf{k})\otimes\mathbf{b}_-(\mathbf{k}).$$

The scalar amplitudes $u_\pm^{(\delta)}$ satisfy the Liouville equations:

$$\begin{cases} \partial_t u_\pm^\delta + \nabla_{\mathbf{k}}H_\pm^\delta\cdot\nabla_{\mathbf{x}}u_\pm^\delta - \nabla_{\mathbf{x}}H_\pm^\delta\cdot\nabla_{\mathbf{k}}u_\pm^\delta = 0, \\ u_\pm^\delta(0,\mathbf{x},\mathbf{k}) = u_\pm^0(\mathbf{x},\mathbf{k}). \end{cases} \tag{5.11}$$

These equations are of the form (2.8), written in the macroscopic variables, with the Hamiltonian given by (5.7).

One may obtain an L^2-error estimate for this convergence when a mixture of states is introduced, as in (5.9). In order to make the scale separation $\varepsilon \ll \delta \ll 1$ precise we define the set

$$\mathcal{K}_\mu := \left\{(\varepsilon,\delta):\ \ |\ln\varepsilon|^{-2/3+\mu} \le \delta \le 1\right\}.$$

The parameter μ is a fixed number in the interval $(0, 2/3)$. The following proposition has been proved in Theorem 3.2 of [1], using straightforward if tedious asymptotic expansions.

Proposition 5.1 *Let the acoustic speed $c_\delta(\mathbf{x})$ be of the form (5.2) and such that the Hamiltonian $H_\delta(\mathbf{x})$ given by (5.7) satisfies assumptions (2.3). We assume that the Wigner transform W_ε^δ satisfies*

$$W_\varepsilon^\delta(0,\mathbf{x},\mathbf{k}) \to W_0(\mathbf{x},\mathbf{k}) \text{ strongly in } L^2(\mathbb{R}^d\times\mathbb{R}^d) \text{ as } \mathcal{K}_\mu \ni (\varepsilon,\delta)\to 0. \tag{5.12}$$

We also assume that the limit $W_0 \in C_c^2(\mathbb{R}^d \times \mathbb{R}^d)$ with a support that satisfies

$$\textit{supp}\ W_0(\mathbf{x},\mathbf{k}) \subseteq \mathcal{A}(M) \tag{5.13}$$

for some $M > 0$. Moreover, we assume that the initial limit Wigner transform W_0 is of the form

$$W_0(\mathbf{x},\mathbf{k}) = \sum_{q=\pm} u_q^0(\mathbf{x},\mathbf{k})\Pi_q(\mathbf{k}), \quad \Pi_q(\mathbf{k}) = \mathbf{b}_q(\mathbf{k}) \otimes \mathbf{b}_q(\mathbf{k}). \tag{5.14}$$

Let $U^\delta(t,\mathbf{x},\mathbf{k}) = \sum_{p=\pm} u_p^\delta(t,\mathbf{x},\mathbf{k})\Pi_p(\mathbf{k})$, where the functions u_p^δ satisfy the Liouville equations (5.11). Then there exists a constant $C_1 > 0$ that is independent of δ so that

$$\|W_\varepsilon^\delta(t,\mathbf{x},\mathbf{k}) - U^\delta(t,\mathbf{x},\mathbf{k})\|_2 \le C(\delta)\left(\varepsilon\|W_0\|_{H^2} e^{C_1 t/\delta^{3/2}} + \varepsilon^2 \|W_0\|_{H^3} e^{C_1 t/\delta^{3/2}}\right) + \|W_\varepsilon^\delta(0) - W_0\|_2, \tag{5.15}$$

where $C(\delta)$ is a rational function of δ with deterministic coefficients that may depend on the constant $M > 0$ in the bound (5.13) on the support of W_0.

The Liouville equations (5.11) are of the form (2.8). Therefore, one may pass to the limit $\delta \to 0$ in (5.11) using Theorem 2.1 and conclude that $\mathbb{E}u_\pm^\delta$ converge to the respective solutions of

$$\frac{\partial \bar{u}_\pm}{\partial t} = \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(|\mathbf{k}|^2 D_{mn}(\hat{\mathbf{k}})\frac{\partial \bar{u}_\pm}{\partial k_n}\right) \pm c_0 \hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} \bar{u}_\pm \tag{5.16}$$

with the initial conditions as in (5.11). Here the diffusion matrix $D(\hat{\mathbf{k}}) = [D_{mn}(\hat{\mathbf{k}})]$ is given by

$$D_{mn}(\hat{\mathbf{k}}) = -\frac{1}{2}\int_{-\infty}^{\infty} \frac{\partial^2 R(c_0 s\hat{\mathbf{k}})}{\partial x_n \partial x_m} ds, \tag{5.17}$$

where $R(\mathbf{x})$ is the correlation function of the random field $c_1(\mathbf{x})$: $\mathbb{E}\left[c_1(\mathbf{z})c_1(\mathbf{x}+\mathbf{z})\right] = R(\mathbf{x})$. Furthermore, it follows from Theorem 2.7 that there exists $\alpha_0 > 0$ so that solutions of (5.11) with the initial data of the form $u_\pm^\delta(0,\mathbf{x},\mathbf{k}) = u_\pm^0(\delta^\alpha \mathbf{x},\mathbf{k})$ with $0 < \alpha < \alpha_0$, converge in the long time limit to the solutions of the spatial diffusion equation. More precisely, in that case the function $\bar{u}^\delta(t,\mathbf{x},\mathbf{k}) = u_+^\delta(t/\delta^{2\alpha}, \mathbf{x}/\delta^\alpha, \mathbf{k})$ (and similarly for u_-^δ) converges as $\delta \to 0$ to $w(t,\mathbf{x},k)$ – the solution of the spatial diffusion equation

$$\frac{\partial w}{\partial t} = \sum_{m,n=1}^{d} a_{mn}(k)\frac{\partial^2 w}{\partial x_n \partial x_m}, \tag{5.18}$$

$$w(0,\mathbf{x},k) = \bar{u}_+^0(\mathbf{x};k) := \frac{1}{\Gamma_{d-1}}\int_{\mathbb{S}^{d-1}} u_+^0(\mathbf{x},\mathbf{k})d\Omega(\hat{\mathbf{k}}).$$

with the diffusion matrix a_{mn} given by:

$$a_{nm}(k) = \frac{c_0}{\Gamma_{d-1}}\int_{\mathbb{S}^{d-1}} \hat{k}_n \chi_m(\mathbf{k}) d\Omega(\hat{\mathbf{k}}), \tag{5.19}$$

and the functions χ_j above are the mean-zero solutions of

$$\sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(k^2 D_{mn}(\hat{\mathbf{k}})\frac{\partial \chi_j}{\partial k_n}\right) = -c_0 \hat{k}_j. \tag{5.20}$$

Theorems 2.1, 2.5 and 2.7 allow us to make the passage to the limit $\varepsilon, \delta, \gamma \to 0$ rigorous. The assumption that $\varepsilon \ll \delta \ll \gamma$ is formalized as follows. We let

$$\mathcal{K}_{\mu,\rho} := \left\{ (\varepsilon, \delta, \gamma) : \quad \delta \geq |\ln \varepsilon|^{-2/3+\mu} \text{ and } \gamma \geq \delta^{\rho} \right\},$$

with $0 < \mu < 2/3$, $\rho \in (0,1)$. Suppose also that $u_0^{\pm} \in C_c^3(\mathbb{R}_*^{2d})$ and $\operatorname{supp} u_0^{\pm} \subseteq \mathcal{A}(M)$. Let

$$W^0(\mathbf{x}, \mathbf{k}) := u_+^0(\mathbf{x}, \mathbf{k})\mathbf{b}_+(\mathbf{k}) \otimes \mathbf{b}_+(\mathbf{k}) + u_-^0(\mathbf{x}, \mathbf{k})\mathbf{b}_-(\mathbf{k}) \otimes \mathbf{b}_-(\mathbf{k}), \tag{5.21}$$

and

$$W(t, \mathbf{x}, \mathbf{k}) := w_+(t, \mathbf{x}; \mathbf{k})\mathbf{b}_+(\mathbf{k}) \otimes \mathbf{b}_+(\mathbf{k}) + w_-(t, \mathbf{x}; \mathbf{k})\mathbf{b}_-(\mathbf{k}) \otimes \mathbf{b}_-(\mathbf{k}). \tag{5.22}$$

Our main result regarding the diffusion of wave energy can be stated as follows.

Theorem 5.2 *Assume that the dimension $d \geq 3$ and $M \geq 1$ are fixed. Suppose for some $0 < \mu < 2/3$, $\rho \in (0,1)$ we have, with W^0 as in (5.21) and W_ε^δ defined by (5.9)*

$$\int\limits_{\mathbb{R}^{2d}} \left| \mathbb{E} W_\varepsilon^\delta \left(0, \frac{\mathbf{x}}{\gamma}, \mathbf{k} \right) - W^0(\mathbf{x}, \mathbf{k}) \right|^2 d\mathbf{x} d\mathbf{k} \to 0, \text{ as } (\varepsilon, \delta, \gamma) \to 0 \text{ and } (\varepsilon, \delta, \gamma) \in \mathcal{K}_{\mu,\rho}.$$

Then, there exists $\rho_1 \in (0, \rho]$ such that for any $T > T_* > 0$ we have

$$\sup_{t \in [T_*, T]} \int \left| \mathbb{E} W_\varepsilon^\delta \left(\frac{t}{\gamma^2}, \frac{\mathbf{x}}{\gamma}, \mathbf{k} \right) - W(t, \mathbf{x}, \mathbf{k}) \right|^2 d\mathbf{x} d\mathbf{k} \to 0, \text{ as } (\varepsilon, \delta, \gamma) \to 0 \text{ and } (\varepsilon, \delta, \gamma) \in \mathcal{K}_{\mu,\rho_1}.$$

Here $W(t, \mathbf{x}, \mathbf{k})$ is of the form (5.22) with the functions $w_\pm$ that satisfy (5.18) *with the initial data $w_\pm(0, \mathbf{x}, \mathbf{k}) = u_\pm^0(\mathbf{x}, \mathbf{k})$.*

The proof follows immediately from Theorems 2.1, 2.5 and 2.7 as well as Proposition 5.1.

A The proof of Lemma 3.5.

Given $s \geq \sigma > 0$, $\pi \in \mathcal{C}$ we define the linear approximation of the trajectory

$$\boldsymbol{L}(\sigma, s; \pi) := X(\sigma) + (s - \sigma) H_0'(K(\sigma)) \hat{K}(\sigma) \tag{A.1}$$

and for any $v \in [0,1]$ let

$$\boldsymbol{R}(v, \sigma, s; \pi) := (1 - v)\boldsymbol{L}(\sigma, s; \pi) + v X(s). \tag{A.2}$$

The following simple fact can be verified by a direct calculation, see Lemma 5.4 of [1].

Proposition A.1 *Suppose that $s \geq \sigma \geq 0$ and $\pi \in \mathcal{C}(\delta)$. Then,*

$$|X(s) - \boldsymbol{L}(\sigma, s; \pi)| \leq \tilde{D}(2M_\delta)\sqrt{\delta}(s - \sigma) + \int\limits_\sigma^s |H_0'(K(\rho))\hat{K}(\rho) - H_0'(K(\sigma))\hat{K}(\sigma)| d\rho.$$

We obtain from Proposition A.1 for each $s \geq \sigma$ an error for the first-order approximation of the trajectory

$$|\mathbf{z}^{(\delta)}(s) - \boldsymbol{l}^{(\delta)}(\sigma, s)| \leq \tilde{D}(2M_\delta)\sqrt{\delta}(s-\sigma) + \frac{C(s-\sigma)^2}{2\sqrt{\delta}}, \quad \delta \in (0, \delta_*(M)].$$

Here $\boldsymbol{l}^{(\delta)}(\sigma, s) := \mathbf{z}^{(\delta)}(\sigma) + (s-\sigma)\hat{\mathbf{m}}^{(\delta)}(\sigma)$ is the linear approximation between the times σ and s and

$$C := \sup_{\delta\in(0,\delta_*(M)]} (M_\delta h_0^*(M_\delta) + \tilde{h}_0^*(M_\delta))\tilde{D}(2M_\delta).$$

With no loss of generality we may assume that $\mathbf{x} = 0$ and that there exists k such that $t, u \in [t_k^{(p)}, t_{k+1}^{(p)})$. We shall omit the initial condition in the notation of the solution to (3.11). Throughout this argument we use Proposition A.1 with

$$\sigma(s) := s - \delta^{1-\gamma_A} \text{ for some } \gamma_A \in (0, 1/16 \wedge (1-\epsilon_4)), \quad s \in [t, u]. \tag{A.3}$$

The aforementioned proposition proves that for this choice of σ we have

$$|\boldsymbol{L}^{(\delta)}(\sigma, s) - \boldsymbol{y}^{(\delta)}(s)| \leq C_A \delta^{3/2-2\gamma_A}, \quad \forall\, \delta \in (0, 1]. \tag{A.4}$$

Throughout this section we denote $\tilde{\zeta} = \zeta(\boldsymbol{y}^{(\delta)}(t_1), \boldsymbol{l}^{(\delta)}(t_1), \ldots, \boldsymbol{y}^{(\delta)}(t_n), \boldsymbol{l}^{(\delta)}(t_n))$. We assume first that $G \in C^2(\mathbb{R}_*^d)$ and $\|G\|_2 < +\infty$. Note that

$$G(\boldsymbol{l}^{(\delta)}(u)) - G(\boldsymbol{l}^{(\delta)}(t)) = -\frac{1}{\sqrt{\delta}} \sum_{j=1}^{d} \int_t^u \partial_j G(\boldsymbol{l}^{(\delta)}(s)) F_{j,\delta}\left(s, \frac{\boldsymbol{y}^{(\delta)}(s)}{\delta}, \boldsymbol{l}^{(\delta)}(s)\right) ds. \tag{A.5}$$

We can rewrite then (A.5) in the form $I^{(1)} + I^{(2)} + I^{(3)}$, where

$$I^{(1)} := -\frac{1}{\sqrt{\delta}} \sum_{j=1}^{d} \int_t^u \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) F_{j,\delta}\left(s, \frac{\boldsymbol{y}^{(\delta)}(s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) ds,$$

$$I^{(2)} := \frac{1}{\delta} \sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \partial_j G(\boldsymbol{l}^{(\delta)}(\rho)) \partial_{\ell_i} F_{j,\delta}\left(s, \frac{\boldsymbol{y}^{(\delta)}(s)}{\delta}, \boldsymbol{l}^{(\delta)}(\rho)\right) F_{i,\delta}\left(\rho, \frac{\boldsymbol{y}^{(\delta)}(\rho)}{\delta}, \boldsymbol{l}^{(\delta)}(\rho)\right) ds\, d\rho,$$

$$I^{(3)} := \frac{1}{\delta} \sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \partial_{i,j}^2 G(\boldsymbol{l}^{(\delta)}(\rho)) F_{j,\delta}\left(s, \frac{\boldsymbol{y}^{(\delta)}(s)}{\delta}, \boldsymbol{l}^{(\delta)}(\rho)\right) F_{i,\delta}\left(\rho, \frac{\boldsymbol{y}^{(\delta)}(\rho)}{\delta}, \boldsymbol{l}^{(\delta)}(\rho)\right) ds\, d\rho$$

and σ is given by (A.3). Each of these terms will be estimated separately below.

A.1 The term $\mathbb{E}[I^{(1)}\tilde{\zeta}]$

The term $I^{(1)}$ can be rewritten in the form $J^{(1)} + J^{(2)}$, where

$$J^{(1)} := -\frac{1}{\sqrt{\delta}} \sum_{j=1}^{d} \int_t^u \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma, s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) ds,$$

and

$$J^{(2)} := -\frac{1}{\delta^{3/2}} \sum_{i,j=1}^{d} \int_t^u \int_0^1 \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{R}^{(\delta)}(v, \sigma, s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) (y_i^{(\delta)}(s) - L_i^{(\delta)}(\sigma, s))\, ds\, dv, \tag{A.6}$$

where, see (A.1) and (A.2), $\boldsymbol{L}^{(\delta)}(\sigma,s)=\boldsymbol{L}(\sigma,s;\boldsymbol{y}^{(\delta)}(\cdot),\boldsymbol{l}^{(\delta)}(\cdot))$, $\boldsymbol{R}^{(\delta)}(\sigma,s)=\boldsymbol{R}(\sigma,s;\boldsymbol{y}^{(\delta)}(\cdot),\boldsymbol{l}^{(\delta)}(\cdot))$. We use part (i) of Lemma 3.2 to handle the term $\mathbb{E}[J^{(1)}\tilde{\zeta}]$. Let $\tilde{X}_1(\mathbf{x},k)=-\partial_{x_i}H_1(\mathbf{x},k)$, $\tilde{X}_2(\mathbf{x},k)\equiv 1$,

$$Z=\Theta\left(t_k^{(p)},\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))\tilde{\zeta}$$

and $g_1=(\boldsymbol{L}^{(\delta)}(\sigma,s)\delta^{-1},|\boldsymbol{l}^{(\delta)}(\sigma)|)$. Note that g_1 and Z are both $\mathcal{F}_\sigma$ measurable. We need to verify (3.14). Suppose therefore that $Z\neq 0$. For $\rho\in[0,t_{k-1}^{(p)}]$ we have $|\boldsymbol{L}^{(\delta)}(\sigma,s)-\boldsymbol{y}^{(\delta)}(\rho)|\geq(2q)^{-1}$, provided that $C_A\delta^{3/2-2\gamma_A}<1/(2q)$, which holds for sufficiently small δ, since our assumptions on the exponents $\epsilon_2,\epsilon_3,\gamma_A$ (namely that $\epsilon_2,\epsilon_3<1/8$, $\gamma_A<1/8$) guarantee that $\epsilon_2+\epsilon_3<3/4-\gamma_A/2$. For $\rho\in[t_{k-1}^{(p)},\sigma]$ we have

$$\begin{aligned}(\boldsymbol{L}^{(\delta)}(\sigma,s)-\boldsymbol{y}^{(\delta)}(\rho))\cdot\hat{\boldsymbol{l}}^{(\delta)}\left(t_{k-1}^{(p)}\right)&\geq(s-\sigma)H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)\,\hat{\boldsymbol{l}}^{(\delta)}(\sigma)\cdot\hat{\boldsymbol{l}}^{(\delta)}\left(t_{k-1}^{(p)}\right)\\&+\int_\rho^\sigma\left[H_0'(|\boldsymbol{l}^{(\delta)}(\rho_1)|)+\sqrt{\delta}\,\partial_l H_1\left(\frac{\boldsymbol{y}^{(\delta)}(\rho_1)}{\delta},|\boldsymbol{l}^{(\delta)}(\rho_1)|\right)\right]\hat{\boldsymbol{l}}^{(\delta)}(\rho_1)\cdot\hat{\boldsymbol{l}}^{(\delta)}\left(t_{k-1}^{(p)}\right)d\rho_1\\&\stackrel{(3.12)}{\geq}(s-\sigma)h_*(2M_\delta)\left(1-\frac{2}{N}\right)+\left[h_*(2M_\delta)-\sqrt{\delta}\tilde{D}(2M_\delta)\right](s-\rho)\left(1-\frac{2}{N}\right)\\&\geq(s-\sigma)h_*(2M_\delta)\left(1-\frac{2}{N}\right),\end{aligned}\tag{A.7}$$

provided that $\delta\in(0,\delta_0]$ and δ_0 is sufficiently small. We see from (A.7) that (3.14) is satisfied with $r=(1-2/N)\,h_*(2M_\delta)\delta^{1-\gamma_A}$. Using Lemma 3.2 we estimate

$$\begin{aligned}\left|\mathbb{E}[J^{(1)}\tilde{\zeta}]\right|&\leq\frac{\tilde{D}(2M_\delta)}{\sqrt{\delta}}\|G\|_1\mathbb{E}[\tilde{\zeta}]\int_t^u\phi\left(C_A^{(1)}\frac{s-\sigma}{\delta}\right)ds\\&\leq C_A^{(2)}\|G\|_1\mathbb{E}[\tilde{\zeta}]\delta^{-1/2}\phi\left(C_A^{(1)}\delta^{-\gamma_A}\right)(u-t)\leq C_A^{(3)}\|G\|_1\mathbb{E}[\tilde{\zeta}]\delta(u-t)\end{aligned}\tag{A.8}$$

and $C_A^{(3)}$ exists by virtue of assumption (2.4). On the other hand, the term $J^{(2)}$ defined by (A.6) may be written as $J^{(2)}=J_1^{(2)}+J_2^{(2)}$, where

$$J_1^{(2)}:=-\frac{1}{\delta^{3/2}}\sum_{i,j=1}^d\int_t^u\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))\partial_{y_i}F_{j,\delta}\left(s,\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)(y_i^{(\delta)}(s)-L_i^{(\delta)}(\sigma,s))\,ds$$

and

$$\begin{aligned}J_2^{(2)}:=-\frac{1}{\delta^{5/2}}\sum_{i,j,k=1}^d\int_t^u\int_0^1\int_0^1\partial^2_{y_i,y_k}F_{j,\delta}\left(s,\frac{\boldsymbol{R}^{(\delta)}(\theta v,\sigma,s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)v\\\times\,\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))(y_i^{(\delta)}(s)-L_i^{(\delta)}(\sigma,s))(y_k^{(\delta)}(s)-L_k^{(\delta)}(\sigma,s))\,ds\,dv\,d\theta.\end{aligned}\tag{A.9}$$

The term involving $J_2^{(2)}$ may be handled easily with the help of (A.4) and Lemma 3.1. We have then

$$\begin{aligned}|\mathbb{E}[J_2^{(2)}\tilde{\zeta}]|&\leq C_A^{(4)}\tilde{D}(2M_\delta)\mathbb{E}[\tilde{\zeta}]\|G\|_1(u-t)\delta^{-5/2}\delta^{3-4\gamma_A-4(\epsilon_2+\epsilon_3)}T^2\\&\leq C_A^{(5)}\delta^{1/2-4\gamma_A-4(\epsilon_2+\epsilon_3)}T^2(u-t)\mathbb{E}[\tilde{\zeta}]\|G\|_1.\end{aligned}\tag{A.10}$$

In order to estimate the term corresponding to $J_1^{(2)}$ we write $J_1^{(2)} = J_{1,1}^{(2)} + J_{1,2}^{(2)}$, where

$$J_{1,1}^{(2)} := -\frac{1}{\delta^{3/2}} \sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) \times (s-\rho_1) \frac{d}{d\rho_1}\left[H_0'(|\boldsymbol{l}^{(\delta)}(\rho_1)|)\, \hat{l}_i^{(\delta)}(\rho_1)\right] ds\, d\rho_1 \tag{A.11}$$

and

$$J_{1,2}^{(2)} := -\frac{1}{\delta} \sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) \partial_l H_1\left(\frac{\boldsymbol{y}^{(\delta)}(\rho)}{\delta}, |\boldsymbol{l}^{(\delta)}(\rho)|\right) \hat{l}_i^{(\delta)}(\rho)\, ds\, d\rho,$$

with

$$\begin{aligned}\frac{d}{d\rho_1}\left[H_0'(|\boldsymbol{l}^{(\delta)}(\rho_1)|)\, \hat{l}_i^{(\delta)}(\rho_1)\right] &= H_0''(|\boldsymbol{l}^{(\delta)}(\rho_1)|)\, (\hat{\boldsymbol{l}}^{(\delta)}(\rho_1), \frac{d}{d\rho_1}\boldsymbol{l}^{(\delta)}(\rho_1))_{\mathbb{R}^d}\, \hat{l}_i^{(\delta)}(\rho_1) \\ &+ H_0'(|\boldsymbol{l}^{(\delta)}(\rho_1)|) |\boldsymbol{l}^{(\delta)}(\rho_1)|^{-1} \left[\frac{d}{d\rho_1} l_i^{(\delta)}(\rho_1) - (\hat{\boldsymbol{l}}^{(\delta)}(\rho_1), \frac{d}{d\rho_1}\boldsymbol{l}^{(\delta)}(\rho_1))_{\mathbb{R}^d}\, \hat{l}_i^{(\delta)}(\rho_1)\right].\end{aligned} \tag{A.12}$$

We deal with $J_{1,2}^{(2)}$ first. It may be split as $J_{1,2}^{(2)} = J_{1,2,1}^{(2)} + J_{1,2,2}^{(2)} + J_{1,2,3}^{(2)}$, where

$$J_{1,2,1}^{(2)} := -\frac{1}{\delta} \sum_{i,j=1}^{d} \int_t^u (s-\sigma) \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) \times \partial_l H_1\left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,\rho)}{\delta}, |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma)\, ds \tag{A.13}$$

$$J_{1,2,2}^{(2)} := -\frac{1}{\delta^2} \sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \int_0^1 \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) \times (\partial_{y_i}\partial_l H_1)\left(\frac{\boldsymbol{R}^{(\delta)}(v,\sigma,\rho)}{\delta}, |\boldsymbol{l}^{(\delta)}(\rho)|\right) (y_i^{(\delta)}(\rho) - L_i^{(\delta)}(\sigma,\rho))\, \hat{l}_i^{(\delta)}(\rho)\, ds\, d\rho\, dv$$

and

$$J_{1,2,3}^{(2)} := -\frac{1}{\delta} \sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \int_\sigma^\rho \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) \times \frac{d}{d\rho_1}\left[\partial_l H_1\left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, |\boldsymbol{l}^{(\delta)}(\rho_1)|\right) \hat{l}_i^{(\delta)}(\rho_1)\right] ds\, d\rho\, d\rho_1.$$

By virtue of (A.4), definition (3.10) and Lemma 3.1 we obtain easily that

$$|\mathbb{E}[J_{1,2,2}^{(2)}\tilde{\zeta}]| \le C_A^{(6)} \delta^{1/2-(3\gamma_A+2\epsilon_2+2\epsilon_3)} \|G\|_1 T(u-t) \mathbb{E}\tilde{\zeta}. \tag{A.14}$$

The same argument and equality (A.12) also allow us to estimate $|\mathbb{E}[J_{1,2,3}^{(2)}\zeta]|$ by the right hand side of (A.14).

Using Lemma 3.1 and the definition (3.10) we conclude that there exists a constant $C_A^{(7)} > 0$ independent of δ such that

$$\left| \partial_{y_i} F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) - \Theta\left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma,s), \boldsymbol{l}^{(\delta)}(\sigma)\right) \partial^2_{y_i,y_j} H_1\left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, |\boldsymbol{l}^{(\delta)}(\sigma)|\right)\right|$$
$$\le C_A^{(7)} \delta^{1-2(\epsilon_2+\epsilon_3)} T, \quad i,j = 1,\ldots,d.$$

Therefore, we can write

$$\left| \mathbb{E}[J^{(2)}_{1,2,1}\tilde{\zeta}] + \frac{1}{\delta}\sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \mathbb{E}\left[\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \Theta\left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma,s), \boldsymbol{l}^{(\delta)}(\sigma)\right)\right.\right. \tag{A.15}$$
$$\left.\left. \times\ \partial^2_{y_i,y_j} H_1\left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \partial_l H_1\left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,\rho)}{\delta}, |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma)\,\tilde{\zeta}\right] ds\, d\rho \right|$$
$$\le C_A^{(8)} \delta^{1-\gamma_A-2(\epsilon_2+\epsilon_3)} (u-t)\|G\|_1 T \mathbb{E}\tilde{\zeta}.$$

With our choice of the exponents we have $\delta < (2q)^{[1-\gamma_A-2(\epsilon_2+\epsilon_3)]^{-1}}$ for all $\delta \in [0,\delta_0)$ where $\delta_0 > 0$ is sufficiently small. We apply now part (ii) of Lemma 3.2 with

$$Z = \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \Theta\left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma,s), \boldsymbol{l}^{(\delta)}(\sigma)\right) \hat{l}_i^{(\delta)}(\sigma)\,\tilde{\zeta},$$
$$\tilde{X}_1(\mathbf{x},k) := \partial^2_{x_i,x_j} H_1(\mathbf{x},k), \quad \tilde{X}_2(\mathbf{x}) := \partial_k H_1(\mathbf{x},k),$$
$$g_1 := \left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, |\boldsymbol{l}^{(\delta)}(\sigma)|\right), \quad g_2 := \left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,\rho)}{\delta}, |\boldsymbol{l}^{(\delta)}(\sigma)|\right),$$
$$r = C_A^{(9)}(\rho-\sigma), \quad r_1 = C_A^{(9)}(s-\rho).$$

We conclude that

$$\left| \mathbb{E}\left[J^{(2)}_{1,2,1}\zeta\right] + \frac{1}{\delta}\sum_{i,j=1}^{d} \int_t^u \int_\sigma^s \mathbb{E}\left[\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \Theta\left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma,s), \boldsymbol{l}^{(\delta)}(\sigma)\right)\right.\right. \tag{A.16}$$
$$\left.\left. \times \partial^2_{y_i,y_j} R_1\left(\frac{\boldsymbol{L}^{(\delta)}(\sigma,s) - \boldsymbol{L}^{(\delta)}(\sigma,\rho)}{\delta},\ |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma)\tilde{\zeta}\right] ds\, d\rho \right|$$
$$\le C_A^{(8)} \delta^{1-\gamma_A-2(\epsilon_2+\epsilon_3)} (u-t)\|G\|_1 T \mathbb{E}\tilde{\zeta}$$
$$+ \frac{C_A^{(10)}}{\delta}\|G\|_1 \mathbb{E}[\tilde{\zeta}] \int_t^u \int_\sigma^s \phi^{1/2}\left(\frac{C_A^{(9)}(\rho-\sigma)}{2\delta}\right) \phi^{1/2}\left(\frac{C_A^{(9)}(s-\rho)}{2\delta}\right) ds\, d\rho,$$

where

$$R_1(\mathbf{y},k) := \mathbb{E}[H_1(\mathbf{y},k)\partial_k H_1(\mathbf{0},k)], \quad (\mathbf{y},k) \in \mathbb{R}^d \times [0,+\infty). \tag{A.17}$$

We can use assumption (2.4) to estimate the second term on the right hand side of (A.16) e.g. by $C_A^{(11)}\delta(u-t)\|G\|_1\mathbb{E}\tilde{\zeta}$. The second term appearing on the left hand side of (A.16) equals to

$$\sum_{j=1}^{d} \int_t^u \mathbb{E}\left\{\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \Theta\left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma,s), \boldsymbol{l}^{(\delta)}(\sigma)\right)\right. \tag{A.18}$$

$$\times \frac{1}{H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)} \left[-\int_\sigma^s \frac{d}{d\rho} \partial_{y_j} R_1 \left(\frac{s-\rho}{\delta} H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)\, \hat{\boldsymbol{l}}^{(\delta)}(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)| \right) d\rho \right] \tilde{\zeta} \Bigg\} ds$$

and integrating over $d\rho$ we obtain that it equals

$$-\sum_{j=1}^d \int_t^u \mathbb{E} \left\{ \frac{\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))}{H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)} \Theta \left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma, s), \boldsymbol{l}^{(\delta)}(\sigma) \right) \partial_{y_j} R_1 \left(\mathbf{0}, |\boldsymbol{l}^{(\delta)}(\sigma)| \right) \tilde{\zeta} \right\} ds \tag{A.19}$$

$$+\sum_{j=1}^d \int_t^u \mathbb{E} \left\{ \frac{\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))}{H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)} \Theta \left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma, s), \boldsymbol{l}^{(\delta)}(\sigma) \right) \partial_{y_j} R_1 \left(\delta^{-\gamma_A} \hat{\boldsymbol{l}}^{(\delta)}(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)| \right) \tilde{\zeta} \right\} ds.$$

By virtue of (2.5) the second term appearing in (A.19) is bounded e.g. by $C_A^{(12)} \delta (u-t) \|G\|_1 \mathbb{E}\tilde{\zeta}$ for some constant $C_A^{(12)} > 0$, thus we have shown that

$$\left| \mathbb{E}[J_{1,2,1}^{(2)} \zeta] - \sum_{j=1}^d \int_t^u \mathbb{E} \left\{ \frac{\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))}{H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)} \Theta \left(t_k^{(p)}, \boldsymbol{L}^{(\delta)}(\sigma, s), \boldsymbol{l}^{(\delta)}(\sigma) \right) \partial_{y_j} R_1 \left(\mathbf{0}, |\boldsymbol{l}^{(\delta)}(\sigma)| \right) \tilde{\zeta} \right\} ds \right| \tag{A.20}$$

$$\le C_A^{(13)} \delta^{1-\gamma_A - 2(\epsilon_2+\epsilon_3)} (u-t) \|G\|_1 T \mathbb{E}\tilde{\zeta}.$$

Let us consider the term corresponding to $J_{1,1}^{(2)}$, cf. (A.11). Note that according to (A.12) and (3.11) we have $J_{1,1}^{(2)} = J_{1,1,1}^{(2)} + J_{1,1,2}^{(2)}$, where

$$J_{1,1,1}^{(2)} := -\frac{1}{\delta^2} \sum_{i,j=1}^d \int_t^u \int_\sigma^s \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta} \left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma, s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma) \right)$$
$$\times (s-\rho_1) \Gamma_i \left(\rho_1, \frac{\boldsymbol{y}^{(\delta)}(\rho_1)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma) \right) ds\, d\rho_1,$$

with

$$\Gamma_i(\rho, \mathbf{y}, \mathbf{l}) := |\mathbf{l}|^{-1} H_0'(|\mathbf{l}|) \left[\left(\hat{\mathbf{l}}, F_\delta(\rho, \mathbf{y}, \mathbf{l}) \right)_{\mathbb{R}^d} l_i - F_{i,\delta}(\rho, \mathbf{y}, \mathbf{l}) \right] - H_0''(|\mathbf{l}|) \left(\hat{\mathbf{l}}, F_\delta(\rho, \mathbf{y}, \mathbf{l}) \right)_{\mathbb{R}^d} \hat{l}_i,$$

while

$$J_{1,1,2}^{(2)} := -\frac{1}{\delta^2} \sum_{i,j=1}^d \int_t^u \int_\sigma^s \int_\sigma^{\rho_1} \partial_j G(\boldsymbol{l}^{(\delta)}(\sigma)) \partial_{y_i} F_{j,\delta} \left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma, s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma) \right)$$
$$\times (s-\rho_1) \frac{d}{d\rho_2} \Gamma_i \left(\rho_1, \frac{\boldsymbol{y}^{(\delta)}(\rho_1)}{\delta}, \boldsymbol{l}^{(\delta)}(\rho_2) \right) ds\, d\rho_1\, d\rho_2. \tag{A.21}$$

Note that $|\frac{d}{d\rho_2} \Gamma_i| \le C_A^{(14)} \delta^{-1/2}$ for some constant $C_A^{(14)} > 0$. A straightforward computation, using (A.3) and Lemma 3.1, shows that $|\mathbb{E}[J_{1,1,2}^{(2)} \zeta]| \le C_A^{(15)} \delta^{1/2 - (3\gamma_A + 2\epsilon_2 + 2\epsilon_3)} (u-t) \|G\|_1 T \mathbb{E}[\tilde{\zeta}]$. An application of (A.4), in the same fashion as it was done in the calculations concerning the terms $\mathbb{E}[J_{1,2,2}^{(2)} \zeta]$

and $\mathbb{E}[J^{(2)}_{1,2,3}\zeta]$, yields that

$$\left|\mathbb{E}[J^{(2)}_{1,1,1}\zeta] + \frac{1}{\delta^2}\sum_{i,j=1}^{d}\int_t^u\int_\sigma^s (s-\rho_1)\mathbb{E}\left[\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))\partial_{y_i}F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right)\right.\right. \tag{A.22}$$

$$\left.\left.\times\ \Gamma_i\left(\rho_1, \frac{\boldsymbol{L}^{(\delta)}(\sigma,\rho_1)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right)\tilde{\zeta}\right]ds\,d\rho_1\right| \le C_A^{(16)}\delta^{1/2-(4\gamma_A+2\epsilon_2+2\epsilon_3)}(u-t)\|G\|_1 T\mathbb{E}[\tilde{\zeta}].$$

For $j=1,\ldots,d$ we let

$$V_j(\mathbf{y},\mathbf{y}',\mathbf{l}) := \sum_{i,k=1}^{d}\left(H_0''(|\mathbf{l}|)-H_0'(|\mathbf{l}|)\right)\partial^3_{y_i,y_j,y_k}R(\mathbf{y}-\mathbf{y}',|\mathbf{l}|)\hat{l}_i\hat{l}_k$$

$$+\sum_{i=1}^{d}H_0'(|\mathbf{l}|)|\mathbf{l}|^{-1}\partial^3_{y_i,y_i,y_j}R(\mathbf{y}-\mathbf{y}',|\mathbf{l}|),$$

and also

$$\Lambda(t,\mathbf{y},\mathbf{y}',\mathbf{l};\pi) := \Theta(t,\mathbf{y},\mathbf{l};\pi)\Theta(t,\mathbf{y}',\mathbf{l};\pi),\quad t\ge 0,\ \mathbf{y},\mathbf{y}'\in\mathbb{R}^d,\ \mathbf{l}\in\mathbb{R}^d_*,\ \pi\in\mathcal{C}, \tag{A.23}$$

$P := \left(\boldsymbol{L}^{(\delta)}(\sigma,s), \boldsymbol{L}^{(\delta)}(\sigma,\rho_1), \boldsymbol{l}^{(\delta)}(\sigma)\right)$, $P_\delta := \left(\delta^{-1}\boldsymbol{L}^{(\delta)}(\sigma,s), \delta^{-1}\boldsymbol{L}^{(\delta)}(\sigma,\rho_1), \boldsymbol{l}^{(\delta)}(\sigma)\right)$ and

$$\overline{\Theta}(s) := \Theta(s, \boldsymbol{y}^{(\delta)}(s), \boldsymbol{l}^{(\delta)}(s); \boldsymbol{y}^{(\delta)}(\cdot), \boldsymbol{l}^{(\delta)}(\cdot)).$$

Applying Lemma 3.1 and part ii) of Lemma 3.2, as in (A.15) and (A.16), we conclude that the difference between the second term on the left hand side of (A.22) and

$$\frac{1}{\delta^2}\sum_{j=1}^{d}\int_t^u\int_\sigma^s (s-\rho_1)\mathbb{E}\left[\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))\Lambda(\sigma,P)V_j\left(P_\delta\right)\tilde{\zeta}\right]ds\,d\rho_1, \tag{A.24}$$

can be estimated by $C_A^{(17)}\delta^{\gamma_A^{(1)}}(u-t)\|G\|_1\mathbb{E}[\tilde{\zeta}]$ for some $\gamma_A^{(1)}>0$. Using the fact that

$$|\boldsymbol{l}^{(\delta)}(\rho)-\boldsymbol{l}^{(\delta)}(\sigma)| \le C_A^{(22)}\delta^{1/2-\gamma_A},\quad \rho\in[\sigma,s], \tag{A.25}$$

estimate (A.4) and Lemma 3.1 we can argue that

$$\left|\Lambda\left(\sigma,P\right)-\overline{\Theta}^2(s)\right| \le C_A^{(18)}(\delta^{1/2-\gamma_A-\epsilon_1}+\delta^{1/2-2(\gamma_A+\epsilon_2+\epsilon_3)}T).$$

We conclude therefore that the magnitude of the difference between the expression in (A.24) and

$$\frac{1}{\delta^2}\sum_{j=1}^{d}\int_t^u\mathbb{E}\left[\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))\overline{\Theta}^2(s)\left(\int_\sigma^s(s-\rho_1)V_j(P_\delta)\,d\rho_1\right)\tilde{\zeta}\right]ds, \tag{A.26}$$

can be estimated by $C_A^{(19)}\delta^{\gamma_A^{(2)}}(u-t)\|G\|_1T\mathbb{E}[\tilde{\zeta}]$ for some $\gamma_A^{(2)}>0$. Using shorthand notation $Q(\sigma) := H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)\,\hat{\boldsymbol{l}}^{(\delta)}(\sigma)$ we can write the integral from σ to s appearing above as being equal to

$$\frac{1}{\delta^2}\int_{s-\delta^{1-\gamma_A}}^{s}(s-\rho_1)\left[\sum_{i,k=1}^{d}\left(H_0''(|\boldsymbol{l}^{(\delta)}(\sigma)|)-H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)\right)\partial^3_{y_i,y_j,y_k}R\left(\frac{s-\rho_1}{\delta}Q(\sigma),|\boldsymbol{l}^{(\delta)}(\sigma)|\right)\right.$$

$$\left.\times\hat{l}_i^{(\delta)}(\sigma)\hat{l}_k^{(\delta)}(\sigma)+H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)|\boldsymbol{l}^{(\delta)}(\sigma)|^{-1}\sum_{i=1}^{d}\partial^3_{y_i,y_i,y_j}R\left(\frac{s-\rho_1}{\delta}Q(\sigma),|\boldsymbol{l}^{(\delta)}(\sigma)|\right)\right]d\rho_1,$$

which upon the change of variables $\rho_1 := (s-\rho_1)/\delta$ is equal to

$$\int\limits_0^{\delta^{-\gamma_A}} \rho_1 \left[\sum_{i,k=1}^{d} \left(H_0''(|\boldsymbol{l}^{(\delta)}(\sigma)|) - H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|) \right) \partial^3_{y_i,y_j,y_k} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma) \hat{l}_k^{(\delta)}(\sigma) \right.$$
$$\left. + H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)|\boldsymbol{l}^{(\delta)}(\sigma)|^{-1} \sum_{i=1}^{d} \partial^3_{y_i,y_i,y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \right] d\rho_1. \tag{A.27}$$

Using the fact that

$$\sum_{k=1}^{d} \partial^3_{y_i,y_j,y_k} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_k^{(\delta)}(\sigma) = \frac{1}{H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)} \frac{d}{d\rho_1} \left[\partial^2_{y_i,y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \right]$$

we obtain, upon integrating by parts in the first term on the right hand side of (A.27), that this expression equals

$$H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)^{-1} \left(H_0''(|\boldsymbol{l}^{(\delta)}(\sigma)|) - H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|) \right) \sum_{i=1}^{d} \left[\delta^{-\gamma_A} \partial^2_{y_i,y_j} R\left(\delta^{-\gamma_A}\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma) \right.$$
$$\left. - \int\limits_0^{\delta^{-\gamma_A}} \partial^2_{y_i,y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma)\, d\rho_1 \right] \tag{A.28}$$
$$+ H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)|\boldsymbol{l}^{(\delta)}(\sigma)|^{-1} \int\limits_0^{\delta^{-\gamma_A}} \rho_1 \partial^3_{y_i,y_i,y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) d\rho_1. \tag{A.29}$$

Note that $\nabla R(\mathbf{0}) = \mathbf{0}$ and

$$\sum_{i=1}^{d} \partial^2_{y_i,y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma) = \frac{1}{H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)} \frac{d}{d\rho_1} \partial_{y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right).$$

We obtain therefore that the expression in (A.28) equals

$$H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)^{-1} \left(H_0''(|\boldsymbol{l}^{(\delta)}(\sigma)|) - H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|) \right) \left[\sum_{i=1}^{d} \delta^{-\gamma_A} \partial^2_{y_i,y_j} R\left(\delta^{-\gamma_A}\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \hat{l}_i^{(\delta)}(\sigma) \right.$$
$$\left. - H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)^{-1} \partial_{y_j} R\left(\delta^{-\gamma_A}\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) \right] \tag{A.30}$$
$$+ H_0'(|\boldsymbol{l}^{(\delta)}(\sigma)|)|\boldsymbol{l}^{(\delta)}(\sigma)|^{-1} \sum_{i=1}^{d} \int\limits_0^{\delta^{-\gamma_A}} \rho_1 \partial^3_{y_i,y_i,y_j} R\left(\rho_1\, Q(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|\right) d\rho_1.$$

Recalling assumption (2.5) we conclude that the expressions corresponding to the first two terms appearing in (A.30) are of order of magnitude $O(\delta^{\gamma_A^{(3)}})$ for some $\gamma_A^{(3)} > 0$. Summarizing work done

in this section, we have shown that

$$\left|\mathbb{E}\left\{\left[I^{(1)}-\sum_{j=1}^{d}\int_{t}^{u}C_j(\boldsymbol{l}^{(\delta)}(\sigma))\overline{\Theta}^2(s)\partial_j G(\boldsymbol{l}^{(\delta)}(\sigma))\,ds\right]\,\tilde{\zeta}\right\}\right|\le C_A^{(20)}\delta^{\gamma_A^{(4)}}(u-t)\|G\|_1T^2\mathbb{E}\tilde{\zeta} \tag{A.31}$$

for some constants $C_A^{(20)},\gamma_A^{(4)}>0$ and (cf. (A.17))

$$C_j(\mathbf{l}):=E_j(\hat{\mathbf{l}},|\mathbf{l}|)+\frac{\partial_{y_j}R_1\left(\mathbf{0},|\mathbf{l}|\right)}{H_0'(|\mathbf{l}|)},$$

$$E_j(\hat{\mathbf{l}},k):=-\frac{H_0'(k)}{k}\sum_{i=1}^{d}\int_0^{+\infty}\rho_1\partial^3_{y_i,y_i,y_j}R\left(\rho_1\,H_0'(k)\hat{\mathbf{l}},k\right)\,d\rho_1,\qquad j=1,\ldots,d.$$

A.1.1 The terms $\mathbb{E}[I^{(2)}\tilde{\zeta}]$ and $\mathbb{E}[I^{(3)}\tilde{\zeta}]$

The calculations concerning these terms essentially follow the respective steps performed in the previous section so we only highlight their main points. First, we note that the difference between $\mathbb{E}[I^{(2)}\tilde{\zeta}]$ and

$$\frac{1}{\delta}\sum_{i,j=1}^{d}\int_t^u\int_\sigma^s\mathbb{E}\left[\partial_jG(\boldsymbol{l}^{(\delta)}(\sigma))\partial_{\ell_i}F_{j,\delta}\left(s,\frac{\boldsymbol{y}^{(\delta)}(s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)F_{i,\delta}\left(\rho,\frac{\boldsymbol{y}^{(\delta)}(\rho)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)\tilde{\zeta}\right]ds\,d\rho \tag{A.32}$$

is less than, or equal to $C_A^{(21)}\delta^{\gamma_A^{(5)}}(u-t)\|G\|_1\mathbb{E}[\tilde{\zeta}]$, cf. (A.25). Next we note that (A.32) equals

$$\begin{aligned}
&\frac{1}{\delta}\sum_{i,j=1}^{d}\int_t^u\int_\sigma^s\mathbb{E}\left[\partial_jG(\boldsymbol{l}^{(\delta)}(\sigma))\partial_{\ell_i}F_{j,\delta}\left(s,\frac{\boldsymbol{L}^{(\delta)}(\sigma,s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)F_{i,\delta}\left(\rho,\frac{\boldsymbol{L}^{(\delta)}(\sigma,\rho)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)\tilde{\zeta}\right]ds\,d\rho\\
&+\frac{1}{\delta^2}\sum_{i,j,k=1}^{d}\int_t^u\int_\sigma^s\int_0^1\mathbb{E}\left[\partial_jG(\boldsymbol{l}^{(\delta)}(\sigma))\partial_{\ell_i}\partial_{y_k}F_{j,\delta}\left(s,\frac{\boldsymbol{R}^{(\delta)}(v,\sigma,s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)\right.\\
&\left.\times F_{i,\delta}\left(\rho,\frac{\boldsymbol{L}^{(\delta)}(\sigma,\rho)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)(y_k^{(\delta)}(s)-L_k^{(\delta)}(\sigma,s))\tilde{\zeta}\right]ds\,d\rho\,dv\\
&+\frac{1}{\delta^2}\sum_{i,j,k=1}^{d}\int_t^u\int_\sigma^s\int_0^1\mathbb{E}\left[\partial_jG(\boldsymbol{l}^{(\delta)}(\sigma))\partial_{\ell_i}F_{j,\delta}\left(s,\frac{\boldsymbol{y}^{(\delta)}(s)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)\right.\\
&\left.\times\,\partial_{y_k}F_{i,\delta}\left(\rho,\frac{\boldsymbol{R}^{(\delta)}(v,\sigma,\rho)}{\delta},\boldsymbol{l}^{(\delta)}(\sigma)\right)(y_k^{(\delta)}(\rho)-L_k^{(\delta)}(\sigma,\rho))\tilde{\zeta}\right]ds\,d\rho\,dv.
\end{aligned} \tag{A.33}$$

A straightforward argument using Lemma 3.1 and (A.4) shows that both the second and third terms of (A.33) can be estimated by $C_A^{(23)}\delta^{1/2-(3\gamma_A+2\epsilon_2+2\epsilon_3)}(u-t)\|G\|_1T^2\mathbb{E}[\tilde{\zeta}]$. The first term, on the other hand, can be handled with the help of part ii) of Lemma 3.2 in the same fashion as we have dealt with the term $J_{1,2,1}^{(2)}$, given by (A.13) of Section A.1, and we obtain that

$$\left|\mathbb{E}\left\{\left[I^{(2)}-\sum_{j=1}^{d}\int_t^u\left(D_j(|\boldsymbol{l}^{(\delta)}(\sigma)|)\overline{\Theta}^2(s)+J_j(s;\boldsymbol{y}^{(\delta)}(\cdot),\boldsymbol{l}^{(\delta)}(\cdot))\overline{\Theta}(s)\right)\partial_jG(\boldsymbol{l}^{(\delta)}(\sigma))\,ds\right]\tilde{\zeta}\right\}\right| \tag{A.34}$$

$$\leq C_A^{(24)} \delta^{\gamma_A^{(6)}} (u-t) \|G\|_1 T \mathbb{E}[\tilde{\zeta}].$$

Here

$$D_j(l) := \frac{\partial_{y_j} R_2(\mathbf{0}, l)}{H_0'(l)}, \qquad R_2(\mathbf{y}, l) := \mathbb{E}[\partial_l H_1(\mathbf{y}, l) H_1(\mathbf{0}, l)], \tag{A.35}$$

$$J_j(s; \boldsymbol{y}^{(\delta)}(\cdot), \boldsymbol{l}^{(\delta)}(\cdot)) := -\sum_{i=1}^{d} \overline{\Theta}_i(s) D_{i,j}(\hat{\boldsymbol{l}}^{(\delta)}(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|),$$

$$\overline{\Theta}_i(s) := \partial_{l_i} \Theta(s, \boldsymbol{y}^{(\delta)}(s), \boldsymbol{l}^{(\delta)}(s); \boldsymbol{y}^{(\delta)}(\cdot), \boldsymbol{l}^{(\delta)}(\cdot)).$$

Finally, concerning the limit of $\mathbb{E}[I^{(3)}\tilde{\zeta}]$, another application of (A.4) yields

$$\left| \mathbb{E}[I^{(3)}\tilde{\zeta}] - \mathcal{I} \right| \leq C_A^{(25)} \delta^{\gamma_A^{(7)}} (u-t) \|G\|_1 \mathbb{E}[\tilde{\zeta}], \tag{A.36}$$

where

$$\mathcal{I} := \frac{1}{\delta} \int_t^u \int_\sigma^s \mathbb{E}\left\{ \partial_{i,j}^2 G(\boldsymbol{l}^{(\delta)}(\sigma)) F_{j,\delta}\left(s, \frac{\boldsymbol{L}^{(\delta)}(\sigma, s)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) F_{i,\delta}\left(\rho, \frac{\boldsymbol{L}^{(\delta)}(\sigma, \rho)}{\delta}, \boldsymbol{l}^{(\delta)}(\sigma)\right) \tilde{\zeta} \right\} ds\, d\rho.$$

Then, we can use part ii) of Lemma 3.2 in order to obtain

$$\left| \mathcal{I} - \sum_{i,j=1}^{d} \int_t^u D_{i,j}(\hat{\boldsymbol{l}}^{(\delta)}(\sigma), |\boldsymbol{l}^{(\delta)}(\sigma)|) \overline{\Theta}^2(s) \partial_{i,j}^2 G(\boldsymbol{l}^{(\delta)}(\sigma))\, ds \right| \leq C_A^{(26)} \delta^{\gamma_A^{(8)}} (u-t) \|G\|_2 T \mathbb{E}[\tilde{\zeta}]. \tag{A.37}$$

Next we replace the argument σ, in formulas (A.31), (A.34) and (A.36), by s. This can be done thanks to estimate (A.25) and the assumption on the regularity of the random field $H_1(\cdot,\cdot)$. In order to make this approximation work we will be forced to use the third derivative of $G(\cdot)$.

Finally (cf. (A.17), (A.35)) note that

$$\nabla_{\mathbf{y}} R_1(\mathbf{0}, l) + \nabla_{\mathbf{y}} R_2(\mathbf{0}, l) = \nabla_{\mathbf{y}\,|\,\mathbf{y}=\mathbf{0}} \mathbb{E}[\partial_l H_1(\mathbf{y}, l) H_1(\mathbf{y}, l)] = \mathbf{0}.$$

Hence we conclude that the assertion of the lemma holds for any function $G \in C^3(\mathbb{R}_*^d)$ satisfying $\|G\|_3 < +\infty$. Generalization to an arbitrary $G \in C_b^{1,1,3}([0, +\infty) \times \mathbb{R}_*^{2d})$ is fairly standard. Let r_0 be any integer and consider $s_k := t + k r_0^{-1}(u-t)$, $k = 0, \ldots, r_0$. Then

$$\mathbb{E}\left\{ [G(u, \boldsymbol{y}^{(\delta)}(u), \boldsymbol{l}^{(\delta)}(u)) - G(t, \boldsymbol{y}^{(\delta)}(t), \boldsymbol{l}^{(\delta)}(t))] \tilde{\zeta} \right\} \tag{A.38}$$

$$= \sum_{k=0}^{r_0-1} \mathbb{E}\left\{ [G(s_{k+1}, \boldsymbol{y}^{(\delta)}(s_{k+1}), \boldsymbol{l}^{(\delta)}(s_{k+1})) - G(s_k, \boldsymbol{y}^{(\delta)}(s_k), \boldsymbol{l}^{(\delta)}(s_k))] \tilde{\zeta} \right\}.$$

$$= \sum_{k=0}^{r_0-1} \mathbb{E}\left\{ [G(s_k, \boldsymbol{y}^{(\delta)}(s_k), \boldsymbol{l}^{(\delta)}(s_{k+1})) - G(s_k, \boldsymbol{y}^{(\delta)}(s_k), \boldsymbol{l}^{(\delta)}(s_k))] \tilde{\zeta} \right\}$$

$$+ \sum_{k=0}^{r_0-1} \mathbb{E}\left\{ [G(s_{k+1}, \boldsymbol{y}^{(\delta)}(s_{k+1}), \boldsymbol{l}^{(\delta)}(s_k)) - G(s_k, \boldsymbol{y}^{(\delta)}(s_k), \boldsymbol{l}^{(\delta)}(s_k))] \tilde{\zeta} \right\}$$

Using the already proven part of the lemma we obtain

$$\left| \sum_{k=0}^{r_0-1} \mathbb{E}\left\{ [\widehat{N}_{s_{k+1}}(G(s_k, \boldsymbol{y}^{(\delta)}(s_k), \cdot)) - \widehat{N}_{s_k}(G(s_k, \boldsymbol{y}^{(\delta)}(s_k), \cdot))] \tilde{\zeta} \right\} \right| \tag{A.39}$$

$$\leq C_A^{(27)}\delta^{\gamma_A^{(9)}}(u-t)\|G\|_{1,1,3}T^2\mathbb{E}\tilde{\zeta}.$$

On the other hand, the second term on the right hand side of (A.38) equals

$$\sum_{k=0}^{r_0-1}\mathbb{E}\left\{\int\limits_{s_k}^{s_{k+1}}\left\{\partial_\rho+\left[H_0'(|\boldsymbol{l}^{(\delta)}(\rho)|)+\sqrt{\delta}\partial_l H_1\left(\frac{\boldsymbol{y}^{(\delta)}(\rho)}{\delta},|\boldsymbol{l}^{(\delta)}(\rho)|\right)\right]\hat{\boldsymbol{l}}^{(\delta)}(\rho)\cdot\nabla_{\mathbf{y}}\right\}\right. \tag{A.40}$$

$$\left.\times\, G(\rho,\boldsymbol{y}^{(\delta)}(\rho),\boldsymbol{l}^{(\delta)}(s_k))\tilde{\zeta}\,d\rho\right\}$$

The conclusion of the lemma for an arbitrary function $G\in C_b^{1,1,3}([0,+\infty)\times\mathbb{R}_*^{2d})$ is an easy consequence of (A.38)–(A.40) upon passing to the limit with $r_0\to+\infty$. $\square$

References

[1] G. Bal, T. Komorowski and L. Ryzhik, Self-averaging of Wigner Transforms in Random Media, Comm. Math. Phys., **242**, 2003, 81–135.

[2] D. Dürr, S. Goldstein and J. Lebowitz, Asymptotic motion of a classical particle in a random potential in two dimensions: Landau model, Comm. Math. Phys., **113**, 1987, 209–230.

[3] L. Erdös and H. T. Yau, Linear Boltzmann equation as the weak coupling limit of a random Schrödinger Equation, Comm. Pure Appl. Math., **53**, 2000, 667–735.

[4] L. Erdös, M. Salmhofer and H.T. Yau, in preparation, 2004.

[5] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud, *Homogenization limits and Wigner transforms*, Comm. Pure Appl. Math., **50**, 1997, 323–380.

[6] I.I. Gikhman and A.V. Skorochod, *Theory of stochastic processes, vol. 3*, Springer Verlag, 1974, Berlin.

[7] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Springer Verlag, 1998, Berlin.

[8] H. Kesten and G. Papanicolaou. A Limit Theorem For Turbulent Diffusion, Comm. Math. Phys., **65**, 1979, 97–128.

[9] H. Kesten, G. C. Papanicolaou, A Limit Theorem for Stochastic Acceleration, Comm. Math. Phys., **78**, 1980, 19–63.

[10] S. Kusuoka and D. Stroock, Appliactions of the Malliavin calculus, Part II, J. Fac. Sci. Univ. Tokyo, Sect. IA, Math, **32**, 1985, 1–76.

[11] P.-L. Lions and T. Paul, *Sur les mesures de Wigner*, Rev. Mat. Iberoamericana, **9**, 1993, 553–618.

[12] L. Ryzhik, G. Papanicolaou, J. Keller, Transport equations for elastic and other waves in random media, Wave Motion, **24**, 1996, 327-370.

[13] D. Strook, *An Introduction to the Analysis of Paths on a Riemannian Manifold*, Math. Surv. and Monographs v. **74**, 2000.

[14] D. Strook, S. R. S. Varadhan *Multidimensional Diffusion Processes*, Berlin, Heidelberg, New York: Springer-Verlag, 1979.

[15] M. Taylor, *Partial differential equations, vol. 2*, New York, Springer-Verlag, 1996.